Department of Computer Science and Automation

On data-driven systems analyzing, supporting and enhancing users' interaction and experience

Ph.D. Thesis

Candidate
Juan Cruz Benito

Supervisors
Roberto Therón Sánchez, Ph.D.
Francisco José García Peñalvo, Ph.D.

July 2018

Roberto Therón Sánchez, Ph.D., Associate Professor in the Department of Computer Science and Automation at the University of Salamanca (Spain) and Francisco José García Peñalvo, Ph.D., Full Professor in the same department and university, as supervisors of the Ph.D. thesis entitled "On data-driven systems analyzing, supporting and enhancing users' interaction and experience" developed by Juan Cruz Benito

Hereby declare that

This Ph.D. Thesis presents enough merits (theoretical and practical ones) evaluated through the proper assessment, publications and original proposals to be presented and defended publicly.

Salamanca, 11 July 2018.

Roberto Therón Sánchez, Ph.D.	Francisco José García Peñalvo, Ph.D.
University of Salamanca	University of Salamanca

Department of Computer Science and Automation

On data-driven systems analyzing, supporting and enhancing users' interaction and experience

Ph.D. Thesis

Candidate
Juan Cruz Benito

Supervisors
Roberto Therón Sánchez, Ph.D.
Francisco José García Peñalvo, Ph.D.

July 2018

Acknowledgments

First of all, I would like to thank my family —mom, dad, wife, grandpa, grandma, other members of my family— and to my friends for the support received these years in the good moments as well as in the less-good moments. I thank my Ph.D. supervisors for everything they taught me these years and to my colleagues at the GRIAL Research Group for all we learned together. I acknowledge and thank all the people who have collaborated with me during this Ph.D.: the fantastic people from USALPHARMA project, the talented people involved in the Spanish Observatory for University Employability and Employment (OEEU) as well as to my new (and awesome) colleagues from IBM Research. It has been a pleasure to share this time with all of you, I never imagined such kind of outstanding mates in this travel.

Finally, I thank the University of Salamanca —which I (will) proudly consider my alma mater— and to the Junta de Castilla y León for funding my research and Ph.D. Thesis during these years in very different ways.

I would like to dedicate this research work to those born in the "You can't" school. Here is another one who finally could.

Abstract

The research areas of Human-Computer Interaction and Software Architectures have been traditionally treated separately, but in the literature, many authors made efforts to merge them to build better software systems. One of the common gaps between software engineering and usability is the lack of strategies to apply usability principles in the initial design of software architectures. Including these principles since the early phases of software design would help to avoid later architectural changes to include user experience requirements. The combination of both fields (software architectures and Human-Computer Interaction) would contribute to building better interactive software that should include the best from both the systems and user-centered designs. In that combination, the software architectures should enclose the fundamental structure and ideas of the system to offer the desired quality based on sound design decisions.

Moreover, the information kept within a system is an opportunity to extract knowledge about the system itself, its components, the software included, the users or the interaction occurring inside. The knowledge gained from the information generated in a software environment can be used to improve the system itself, its software, the users' experience, and the results. So, the combination of the areas of Knowledge Discovery and Human-Computer Interaction offers ideal conditions to address Human-Computer-Interaction-related challenges. The Human-Computer Interaction focuses on human intelligence, the Knowledge Discovery in computational intelligence, and the combination of both can raise the support of human intelligence with machine intelligence to discover new insights in a world crowded of data.

This Ph.D. Thesis deals with these kinds of challenges: how approaches like data-driven software architectures (using Knowledge Discovery techniques) can help to improve the users' interaction and experience within an interactive system. Specifically, it deals with how to improve the human-computer interaction processes of different kind of stakeholders to improve different aspects such as the user experience or the easiness to accomplish a specific task.

Several research actions and experiments support this investigation. These research actions included performing a systematic literature review and mapping of the literature that was aimed at finding how the software architectures in the literature have been used to support, analyze or enhance the human-computer interaction. Also, the actions included work on four different research scenarios that presented common challenges in the Human-Computer Interaction knowledge area. The case studies that fit into the scenarios selected were chosen based on the Human-Computer Interaction challenges they present, and on the authors' accessibility to them. The four case studies were: an educational laboratory virtual world, a Massive Open Online Course and the social networks where the students discuss and learn, a system that includes very large web forms, and an environment where programmers develop code in the context of quantum computing. The development of the experiences involved the review of more than 2700 papers (only in the literature review phase), the analysis of the interaction of 6000 users in four different contexts or the analysis of 500,000 quantum computing programs.

As outcomes from the experiences, some solutions are presented regarding the minimal software artifacts to include in software architectures, the behavior they should exhibit, the features desired in the extended software architecture, some analytic workflows and approaches to use, or the different kinds of feedback needed to reinforce the users' interaction and experience.

The results achieved led to the conclusion that, despite this is not a standard practice in the literature, the software environments should embrace Knowledge Discovery and data-driven principles to analyze and respond appropriately to the users' needs and improve or support the interaction. To adopt Knowledge Discovery and data-driven principles, the software environments need to extend their software architectures to cover also the challenges related to Human-Computer Interaction. Finally, to tackle the current challenges related to the users' interaction and experience and aiming to automate the software response to users' actions, desires, and behaviors, the interactive systems should also include intelligent behaviors through embracing the Artificial Intelligence procedures and techniques.

Keywords: Human-Computer Interaction, Software Architectures, Knowledge Discovery in Databases, Data-driven, Data, User Experience, Interaction, Users' performance, Artificial Intelligence, Machine Learning, Deep Learning, Virtual Worlds, Massive Open Online Courses, Web forms, Quantum Computing, Programming.

Resumen

Las áreas de investigación sobre la Interacción Persona-Ordenador y las Arquitecturas *Software* han sido tratadas tradicionalmente por separado. A lo largo de los años, muchos autores se han esforzado en unir ambos campos para construir mejores sistemas *software*. Una de las brechas comunes entre la ingeniería del *software* y la usabilidad es la falta de estrategias para aplicar principios de usabilidad en una arquitectura *software* desde las fases de diseño inicial. La inclusión de estos principios desde las fases tempranas del diseño del *software* pueden ayudar a remitir cambios arquitectónicos posteriores que buscan incluir requisitos relacionados con la experiencia de usuario. La mezcla de ambas áreas (la Interacción Persona-ordenador y las arquitecturas software) podría contribuir a construir mejor *software* interactivo que incluyan lo mejor de los sistemas informáticos y de los diseños centrados en el usuario. En esa combinación, las arquitecturas *software* deberían contener una estructura fundamental y las ideas básicas del sistema para ofrecer la calidad deseada en base a unas decisiones de diseño adecuadas.

Por otro lado, la información almacenada en cualquier sistema informático representa un oportunidad de extraer conocimiento sobre el sistema en sí mismo, sus componentes, el *software* que incluye, los usuarios, o la interacción que ocurre internamente entre todos los actores que participan en el mismo. El conocimiento obtenido de la información generada en un entorno software puede usarse para mejorar el sistema en sí, su *software*, la experiencia de los usuarios y los resultados del mismo. Por tanto, la combinación de las áreas de Descubrimiento de Conocimiento y la Interacción Persona-Ordenador ofrece unas condiciones ideales para tratar con los retos que supone la Interacción Persona-Ordenador. La Interacción Persona-Ordenador se centra en la inteligencia humana, mientras que el Descubrimiento de Conocimiento se centra en la inteligencia computacional; la combinación de ambas puede ayudar a soportar la inteligencia humana con la inteligencia computacional para descubrir nuevas evidencias y resultados en un mundo repleto de datos.

Esta Tesis Doctoral trata con ese tipo de retos: cómo aproximaciones como las arquitecturas *software* guiadas por datos (usando técnicas de Descubrimiento de Conocimiento) pueden ayudar a mejorar la interacción de los usuarios y su experiencia dentro de un sistema interactivo. Específicamente, trata sobre cómo mejorar los procesos de interacción persona-ordenador de distintos tipos de usuarios y actores para mejorar distintos aspectos como la experiencia de los usuarios o la facilidad para completar una tarea concreta.

La investigación que se presenta está soportada por diversas acciones de investigación y experimentos. Entre las acciones de investigación se incluye una revisión y mapeo sistemático de la literatura que pretende encontrar cómo se han usado las arquitecturas *software* en la literatura para soportar, analizar o mejorar la interacción entre personas y ordenadores. Estas

acciones también incluyen el trabajo en cuatro escenarios de investigación distintos que presentan retos comunes en el área de conocimiento de la Interacción Persona-Ordenador. Los casos de estudio que encajan en cada uno de estos escenarios han sido seleccionados por los retos que presentan en relación a la Interacción Persona-Ordenador y por la accesibilidad del autor a los mismos. Los cuatro casos de estudio fueron: un laboratorio educativo dentro de un mundo virtual, un Curso On-line Masivo y Abierto y las redes sociales usadas en relación por los estudiantes para discutir y aprender, un sistema *software* que incluye formularios web muy extensos y un entorno donde programadores desarrollan código en el ámbito de la computación cuántica. El desarrollo de estas experiencias ha requerido la revisión de más de 2700 artículos científicos (solo durante la fase de revisión), el análisis de la interacción de 6000 usuarios entre los cuatro contextos distintos o el análisis de 500000 programas que emplean código relacionado con la computación cuántica.

Como resultados de estas experiencia, se presentan diversas soluciones relacionadas con los artefactos software mínimos que se necesitan en una arquitectura *software* que incluya el soporte, análisis o mejora de la interacción entre personas y ordenadores, el comportamiento que deben tener, las características deseadas en dichas arquitecturas extendidas, algunos flujos de trabajo y aproximaciones para el análisis, o los distintos tipos de refuerzo que se puede proporcionar a los usuarios para mejorar su interacción y experiencia.

Los resultados obtenidos permiten concluir que, aunque no es una práctica habitual en la literatura, los entornos *software* deben emplear el Descubrimiento de Conocimiento y los principios de los sistemas guiados por datos para analizar y responder apropiadamente a las acciones, deseos y comportamientos de los usuarios y para mejorar o soportar su interacción. Al adoptar el Descubrimiento de Conocimiento y los principios basados en datos, los sistemas *software* deben extender sus arquitecturas para poder afrontar los retos relacionados con la Interacción Persona-Ordenador. Finalmente, para ser capaces de responder a los problemas actuales en relación a la interacción y experiencia de los usuarios, e intentando automatizar la respuesta del *software* a los deseos, acciones y comportamientos de los usuarios, los sistemas interactivos deben adoptar comportamientos inteligentes a través de los procedimientos y técnicas relacionadas con la Inteligencia Artificial.

Palabras clave: Interacción Persona-Ordenador, Arquitecturas Software, Descubrimiento de conocimiento en bases de datos, Sistemas guiados por datos, Datos, Experiencia de Usuario, Interacción, Rendimiento de usuarios, Inteligencia Artificial, Aprendizaje máquina, Aprendizaje profundo, Mundos Virtuales, Cursos On-line Masivos y Abiertos, Formularios Web, Computación cuántica, Programación.

Table of Contents

Table of Figures

Table of Tables

1 Introduction

Software Architectures are not a new thing in Computer Sciences. Many authors in the literature have defined and pushed forward the study of this concept. In this work, the definitions provided in 1998 by Kruchten [1] and in 2012 by Bass, Clements, and Kazman [2] will be highlighted. According to Kruchten, a software architecture is "[...] the set of significant decisions about the organization of a software system, the selection of the structural elements and their interfaces by which the system is composed, together with their behavior as specified in the collaborations among those elements, the composition of these structural and behavioral elements into progressively larger subsystems, and the architectural style that guides this organization —these elements and their interfaces, their collaborations, and their composition—" [1]. Bass, Clements, and Kazman provide a simpler definition: "the Software Architecture of a system is the set of structures needed to reason about the system, which comprises the software elements, relations among them, and properties of both" [2].

Also, the Human-Computer Interaction and Human-Machine Interaction research areas have been developing for a long time (at least for the computer science time-scale). The Association for Computing Machinery (ACM) defines Human-Computer Interaction (HCI) as "a discipline concerned with the design, evaluation, and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them" [3]. Other authors, such as Alan Dix, define HCI as "the study of the way in which computer technology influences human work and activities" [4]. Moreover, Human-Machine Interaction (HMI) has been defined in the past as "how we as humans interact with machines, and we define a machine as any mechanical or electrical device that transmits or modifies energy to perform or assist in the performance of human tasks" [5]. Despite these definitions, the term HMI is currently used in the same way as HCI, with the only exception that it is associated mostly with robotics and industrial applications and machines [3, 5, 6].

Both areas, one typically related to software engineering, and the other to fields such as perception, psychology, design, user research, etc. may seem to be separated, but researchers have made many efforts to merge them to build better systems. Thus, Tran, Ezzedine, Kolski state that "the architectural design of interactive systems is the object of many pieces of research since the eighties" [7]. Following the idea of the combination of both fields, Seffah, Gulliksen and Desmarais argue that: "one of the common communication breakdowns between software engineers and usability professionals is the lack of strategies to inform the early design of software architectures with usability principles, which helps avoid late (and expensive) architectural changes to accommodate user experience requirements" [8]. What is more, in complex environments such as software ecosystems, the user, and his or her human factors are included as part of the whole, considering it as important as any other software component or restriction [9-11]. According to García-Holgado & García-Peñalvo, "the users are components of the ecosystem, they establish information flows with other components, and they are affected if the ecosystem evolves" [9]. It can be considered that the users' interaction with software components in the system are as important as the interaction between the software components. It may result in considering the HCI process/characteristics to be at the same level as the other components or parts of the (eco)system when designing or developing a system.

The combination of the fields of software architectures and HCI would contribute to building better systems that should include the best from both the systems and user-centered designs. As some authors claim [12, 13], software architectures should enclose the fundamental structure and ideas of the system to offer the desired quality based on sound design decisions. This quality could be related to any part or goal of the system, such as structural quality, presentation quality, etc. Seffah [14] for example, addresses the combination of HCI and software architectures in the following manner: "in the field of human-computer interaction (HCI), interactive systems architectures of the 1980s and 1990s such as MVC (Model, View, Controller) [15] and PAC (Presentation, Abstraction, and Control) [16] are based on the principle of separating the core functionality from the UI. The functionality is what the software does and what data it processes, thereby offering behavior that can be exploited to achieve some user needs. The UI defines how the functionality is presented to end-users and how the users interact with it. The underlying assumption is that usability, the ultimate quality factor, is primarily a property of the UI. Therefore, separating the UI from the application's logic makes it easy to modify, adapt, or customize the interface after user testing. Unfortunately, this assumption does not ensure the usability of the system as a whole". It has been proven that system features can have an impact on the usability of the system, even if they are logically independent of the UI and not necessarily visible to the user. For example, Bass observed that even if the presentation of a system was well designed, the usability of a system could be significantly compromised if the underlying architecture and designs did not have the proper provisions for user concerns [17].

Adhering to these ideas presented in the literature, the motivation and starting point of this thesis can be summarized as follows: the software architecture should not only define the technical issues needed to develop and implement functionality, but also to manage and facilitate the interaction between the users and the system.

1.1 Context of this research

Since many years ago, one of the most relevant challenges in computer science and areas-alike is dealing with large, complex, or unstructured data [18-35]. As it may been seen in these references, working with data is a difficult task: it presents issues related on how to save information (in a physical mode), structure the information, adequately analyze the data, extract knowledge from data, make decisions, etc. Despite these issues, information is one of the most relevant aspects and source of power in the society; for that reason, many authors talk about the existence of the Information Society [36, 37] or the Knowledge Society [38-40]. Related to the definition about the software architectures, discussed in the introduction, and considering that all software components, their relationships and the decisions about them made by engineers, it is possible to think that the data related to users in the system and the data that fuels the system itself are a core part to consider within the software architecture [2, 41-43]. The information kept within a system is an opportunity to extract knowledge about the system itself, its components, the software included, the users or the interaction that occurs inside [44]. As well as others do in other research & development fields, the knowledge gained from the information generated in a software environment can be used to improve the system itself, its software, the users' experience, the results achieved using it, etc. Extending the starting point expressed in the introduction section, this thesis is about these kinds of challenges: how approaches like data-driven software architectures can help to improve the users' interaction with a system. Specifically, it deals with how to improve the interaction of different kinds of stakeholders with software systems to improve different aspects like the user experience, the easiness to accomplish a specific task, etc.

This idea of knowledge distilling from information, leads to the introduction of the Knowledge Discovery in Databases (KDD) research area. KDD can be defined as "an automatic, exploratory analysis and modeling of large data repositories and the organized process of identifying valid, novel, useful and understandable patterns from large and complex data sets" [45, 46].

This term and the related research area became popular at the end of the 20th century, with some of the most known core contributions defining the scope of KDD [47-50], the beginning of the KDD Conference [51], and the publication of foundational papers like the one by Fayyad, Piatetsky-Shapiro and Smyth [52]. In the paper, the authors described the KDD process as a set of 9 chained steps:

1. Learning from the application domain: includes understanding relevant previous knowledge, the goals of the application and a certain amount of domain expertise.
2. Creating a target dataset: includes selecting a dataset or focusing on a subset of variables or data samples on which discovery shall be performed.
3. Data cleansing (and preprocessing): includes removing noise or outliers, strategies for handling missing data, etc.
4. Data reduction and projection: includes finding useful features to represent the data, dimensionality reduction, etc.

5. Choosing the function of data mining: includes deciding the purpose and principle of the model for mining algorithms (e.g., summarization, classification, regression, and clustering).
6. Choosing the data mining algorithm: includes selecting method(s) to be used for searching for patterns in the data, such as deciding which models and parameters may be appropriate (e.g., models for categorical data are different from models on vectors over real numbers) and matching a particular data mining method with the criteria of the KDD process.
7. Data mining: searching for patterns of interest in a representational form or a set of such representations, including classification rules or trees, regression, clustering, sequence modeling, dependency and line analysis.
8. Interpretation: includes interpreting the discovered patterns and possibly returning to any of the previous steps, as well as possible visualization of the extracted patterns, removing redundant or irrelevant patterns and translating the useful ones into terms understandable by users.
9. Using discovered knowledge: includes incorporating this knowledge into the performance of the system, taking actions based on the knowledge or documenting it and reporting it to interested parties, as well as checking for, and resolving, potential conflicts with previously believed knowledge.

Nowadays, KDD has evolved to cover new challenges and data analysis paradigms and approaches. The KDD field is now broader than ever: it can be considered that currently it includes the knowledge discovery in almost any kind of information saving system (and not only the classical databases) as well as it covers more methods of knowledge discovery beyond the query of databases or the Data Mining, as areas like the Machine Learning or the Artificial Intelligence are assumed as fields that could involve KDD [44, 53, 54].

Similarly, HCI has also experimented a significant evolution in recent years. Part of this growth is due to new HCI paradigms [55, 56], the use of new design systems [57, 58], the consideration of the user as the center of the systems [59], the evaluation of user's interaction [60, 61], the consideration of elements of human life such as culture, emotion, experience or psychological aspects [62-64], the changes in use context (mobility, ubiquity, etc.), the multiplicity of interaction in conventional systems, the broader application types in technical systems, the normalization of computing as a daily resource to use [65] or the tailorability or adaptability of systems to users [66, 67]. The change is so disruptive that many authors looking at this shift in HCI, and following what scholars did in fields like History or Art, affirmed that the HCI discipline is living in a new era or wave [55, 56, 68, 69].

The ideas related to gaining knowledge and use it to improve the interaction between users and systems have been explored previously (at least in a theoretical way) by other authors [44, 55, 68, 69]. For example, in words of Holzinger [44] "a synergistic combination of methodologies, methods and approaches of two areas —KDD and HCI— offer ideal conditions for addressing these challenges: HCI, with its emphasis on human intelligence, and KDD, dealing with computational intelligence – with the goal of supporting human intelligence with machine intelligence – to discover new, previously unknown insights within the flood of data" [70].

Is the combination of KDD and HCI needed? In the past, some authors discussed this topic [44, 71] reflecting the benefits given by a multidisciplinary view to solving the problems that involve issues from both research areas. Nowadays, with the popularization of concepts and areas like Data Science [72-77], this discussion is not as intense as previously. Most parts of researchers (and companies) are now convinced about the need and advantage given by having a data-driven culture, a data-driven design and exploitation policy of digital products and the benefits given by a multidisciplinary point of view in gathering concepts, theories and approaches from different research fields through a common scaffold like the data [78, 79].

In the past, the benefits of joining humans and computers were summarized by the following old proverb [80] (many times wrongly attributed to Einstein [81]):

- Computers are incredibly fast, accurate, and stupid.
- Human beings are incredibly slow, inaccurate, and brilliant.
- Together they are powerful beyond imagination.

This quote fits perfectly also in some of the paradigms related to the cooperation between humans and machines like HCI or Human-Centered Computing (HCC) [82]. In short, the HCC paradigm overlaps HCI except for the fact that in HCC, not only the user and its context are observed, but the users' intentions and motives are estimated from the observed behavior (using interaction models and other techniques) [56]. In any case, many authors consider HCC part of the HCI [4, 57, 83]. The concept of how the users are observed and their intentions are anticipated, as HCC suggests, —from a system point of view— is illustrated in Figure 1, where three parts are depicted (apart from the user): the front-end, the back-end and the models of interaction. The front-end (or the system's input interface) deals with observing the user and its context, as well as it allows user's interaction [84]. Meanwhile, the back-end is responsible for controlling the different elements to be used in the front-end, analyzes user's interactions and behavior to generate the proper behavior and response of the system. To analyze the users' interaction and model their behaviors, the back-end employs the models of interaction; otherwise, to deal with the observation of users' behavior, the front-end applies the models of interaction to complete the flow between the system and the user [56]. One example of the issues solved by models of interaction are the dialog management models, or the interaction templates designed by User Experience (UX) engineers to respond to the users' needs.

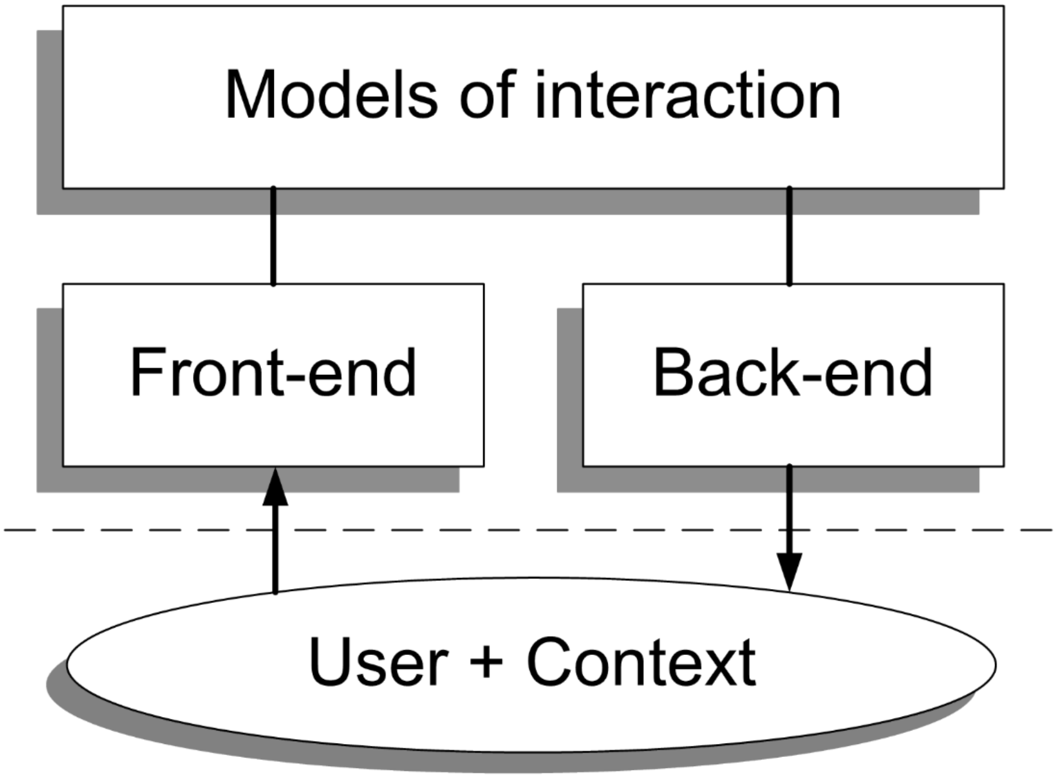


Figure 1. Schematic model of Human-Centered Computing. Taken from [56].

Briefly put, the potential outcomes from the combination of KDD and HCI (or HCC) are related to Figure 1. The workflow needed to analyze, feedback and affect the interaction between users and machines can be highly improved by the application of KDD techniques and principles. In a world where the data are flooding all kinds of systems and the users' interfaces have evolved to be ubiquitous and present many different shapes, the application of knowledge discovery can help systems to evolve, be adaptive and respond better to the users' needs, desires and behaviors. The KDD principles and procedures, the current data science approaches and the advantages of data-driven environments can be a golden opportunity for the HCI field and sub-areas to tackle the current challenges in the area.

1.2 Hypotheses and goals

This academic work is not intended to respond the research questions with a simple answer like 'yes', 'no' or even with heavy-weight words like "it is true because it is statistically significant". It is true that these answers can be convenient to respond some of the specific questions, sub-hypotheses and so on, but, in general, short and rotund answers to the problem posed here, or to the principal —and broad— research questions, are avoided.

Thus, the plan is to dig into the central questions, to later respond more specifically to the different sub-questions and build up, using these responses, a broad answer to the primary questions and discuss them. To develop the research, the ancient strategy *divide and conquer* is used, considering *conquer* as the attempt to respond to a complex question that could include many nuances.

After all those considerations, it is time to pose the essential research questions of this work:

Is it beneficial to follow a data-driven/KDD approach in a system to support, analyze or improve the users' interaction and experience and tackle the HCI-related challenges to these aspects that the interactive systems present?

How a software environment should evolve to respond to the users' needs and improve or support the users' interaction and experience?

How is it possible to do that in a more automated way?

More specifically, the questions could accept responses from all of these sub-questions:

1. What kind of software artifacts does a system require to respond to the users' needs or to be adaptive to users?
2. What features should these kinds of systems have?
3. What kind of software behaviors related to those features should be the common ones?
4. What kind of strategies should these systems include to provide valuable feedback to the users?
5. Could an intelligent system be capable of improving the UX in a significant manner? Would it be adequate to include intelligent features in a system to pursue such UX improvement?

To answer these previous questions, it is planned to explore the different related issues by means of a research using different (and incremental) scenarios. These scenarios are selected following two primary considerations: they are available to the author[1] (research based on the accessibility to the scenarios), and they present relevant human-computer interaction challenges. Many authors have tried to envision essential challenges for the future of HCI [44, 55, 56, 58, 68, 69, 83, 85]. In this research work, the HCI-related challenges presented in [69] (Table 1) have been considered as possible candidates.

[1] Disclaimer: in general, this document uses 'author', 'researcher' or other similar words in the singular form to refer to the people that authored the presented research. Although most of the text is based on papers already published with more than one author, consider it a stylistic license. All the manuscripts used for this thesis have been (co)authored mainly (most of them as the first author) by the Ph.D. candidate, and all the participant authors in the different experiments and experiences have been informed previously and are acknowledged in the document, as the papers are included in Appendixes.

Table 1. Proposed R&D Road Map by Stephanidis et al. (1999) [69]

Promote the Development of Environments of Use	*Support Communities of Users*	*Extend User-Centered Design to Support New Virtualities*	*Establish Suitable Accompanying Measures*
Determine desirable properties of environments of use (e.g., augmented capabilities on user's demand, multimodality, cooperativity, intelligence, adaptation, etc.)	Develop individual and collective intelligence and community knowledge management	Develop suitable foundations for design, by applying, integrating, and extending existing user-centered design methods to facilitate the design of new virtual spaces	Articulate demand for design for all
Develop novel architectures for interactive systems for managing collective experiences of users and nonusers	Develop methodologies for collecting and analyzing requirements and understanding virtual communities	Develop metrics for important quality attributes (e.g., usability, accessibility, adaptation, intelligence, etc.)	Support of the industry
Design architectures for multiple metaphor environments	Provide means to access communitywide information resources	Provide computation support for usability engineering (e.g., computer-supported usability platforms)	Create awareness and knowledge dissemination
Develop multiagent systems and components to support cooperation and collaboration	Develop models to support social interaction among members of online communities	Extend existing requirements for engineering methods to facilitate elicitation of requirements in novel contexts of use and different user groups	Extend technology transfer
Support individualization and user interface adaptation (e.g., adaptability and adaptivity) of environments of use		Promote user involvement and develop protocols for effective user participation in design activities	
		Investigate and provide design recommendations for alternative interaction modalities and their combinations	

Although this paper is quite old (it is from 1999) most of the challenges and the R & D roadmap are still relevant today in the Computer Science field.

Regarding the scenarios, and according to [69] and Table 1, the following prototypical scenarios can be selected and adapted to develop and test the research:

1. Highly interactive scenarios where the users should develop individual activities or tasks based only on their interaction without human aid.
2. Systems including a high number of users collaborating or using concurrently the same resources to solve the same tasks. In this scenario, other stakeholders evaluate the users' performance in solving their tasks.
3. Information-intensive applications where the users could need aid and be engaged in solving appropriately the proposed challenges (with challenge referring to the need for solving a task).
4. Complex systems where the user could not have enough previous knowledge to solve a task efficiently.

Each prototypical scenario has their differences and similarities with the others; also, in general, they are aligned mainly with adaptations of the two first columns of Table 1.

Assuming these scenarios, it is possible to define some objectives for each one of them. The primary goal is related principally to an HCI challenge in the scenario, and the second one is related more to how to integrate the solution related to the HCI challenge in a software system.

- Scenario 1.

    - Primary objective: Collect and analyze individual users' interaction from a highly interactive environment to extract knowledge and evaluate users' performance solving tasks.
    - Secondary objective: test how to take advantage of the knowledge gained, from a software perspective, to provide feedback to the users and improve their results.
- Scenario 2.
    - Primary objective: understand how users collaborate in a large environment to solve (individually) a task shared by all users.
    - Secondary objective: integrate a software solution in a massive environment to automatize the analysis performed in the primary objective.
- Scenario 3.
    - Primary goal: improve the user experience (and engagement) in a complex environment that involves a high amount of information and introduces a high level of friction for users when solving a task. This enhancement of UX seeks to improve the users' performance in the proposed task by reducing the friction.
    - Secondary goal: define an automated data-driven pipeline to analyze the users' interaction and improve the user experience.
- Scenario 4.
    - Primary goal: study and propose a way to aid users in solving a complex task, even when they do not have enough knowledge to solve it.
    - Secondary goal: based on the proposal raised after the primary objective, create software that helps users and could be integrated into different environments related to the problem.

This organization of the scenarios affects the structure of the document and the contents present in each section, as commented in section 1.5.

1.3 Methodology

Given the organization of the research based on the different proposed scenarios and the experiments performed to dig into the research questions, different methodologies and research approaches have been applied in this thesis. This section is aimed at providing an overview of the methodological perspective followed during the research. The details related to the methodological aspects of each experiment and scenarios are presented in each subsection within chapter 1.

The general methodological framework of the research is based on a quantitative perspective. A quantitative methodology is oriented to the development of knowledge. It focuses on the development of the theories through approaching the objectives, hypotheses, and questions. In this case, if a researcher wants to test general hypothesis, designs smaller hypotheses and uses the corresponding data to confirm or refute them [86]. Some typical strategies used in the quantitative methodology are: surveys, experiments and compilation of numerical data that are used to investigate. In the case of this thesis, all of the experiments followed a quantitative approach, since all the experiments use the data-driven perspective, and employ the data to attempt to answer the different research questions and achieve the proposed goals.

According to [86, 87], there are several types of research depending on the purposes of the research, the position of the researcher in front of the research, etc. Since this thesis contains several experiments and experiences, different types of research are employed. Below the most characteristic ones are presented.

Besides using a quantitative approach, this thesis follows a formal process called hypothetic-deductive model [86, 88]. Using the theory from the knowledge area, the author plans an inductive process (from the particular to the general) to be validated by a deductive process (from a general premise to the particular outcomes).

Regarding the approach or level of knowledge that the author wants to achieve, in this thesis both the exploratory and descriptive types of research are performed. Most of the experiments to be presented in the next sections are under the framework of descriptive research (describes the research according to the observations). Also, in a few experiments, especially in the first approaches to the experiment (first sub-experiments) the process can be defined as exploratory research. The exploratory research is typical in under-researched topics.

Considering the researcher's role in the research process, all the experiments performed in this thesis follow a non-experimental design. The experimentation with people (users of systems) has employed in all cases the population of users existing in each system and scenario, so the researcher has not designed previously controlled groups of people, etc., and there is no control of variables prior to the experiment. It is true that during the experiments the author has controlled somehow groups of users, experimental conditions, and so on, but from the research theory, this cannot be considered experimental or quasi-experimental research designs.

The orientation of this thesis is to produce conclusions-oriented research. Using the quantitative framework adopted, the author tried to find some responses to the research questions and hypotheses posed in the previous section.

Regarding the source of the knowledge in the research, this thesis employs two kinds of research: bibliographic and empirical. In chapter 2 the state-of-the-art in software architectures analyzing, supporting or enhancing the human-computer interaction processes is introduced. This kind of research is called bibliographic research and permits to know the state of the art (the theoretical foundation) in a field of knowledge. Chapter 1 introduces the experiments and experiences performed. This kind of research is known as empirical research and bases the outcomes and results on the observation performed.

1.4 Framework

This thesis is, in part, a result of the combination of the work done in multiple projects where the author has been involved during the last five years, as well as the collaborations with people around the world, and author's experiences and personal interests.

First of all, the idea behind this thesis appeared five years ago during the first experiences in research tasks performed within the Master in Intelligent Systems (MSc) at the University of Salamanca. In that MSc, the author began to develop a particular interest in data science, in Human-Computer Interaction, in analyzing user's interaction with virtual environments [89-92], etc. After the Master, the author embraced the opportunity of beginning to work in the GRIAL Research Group [93, 94]. The experience gained while working in this research group has enriched the author's views on the research world profoundly. The experience gained through the participation in different projects during these years, the people met, the conferences organized and attended, and the people that compose the group has been critical in pursuing these results. Also, the different research carried out (in the recent past or currently) by other colleagues at the research group have served as inspiration for the author [10, 11, 95-113].

Apart from that, the different projects and experiences relevant for this research work are explained below. As the reader will observe, most of them are related to Education in very different ways; the projects in which the candidate worked and the personal interest in some areas related to Education and how people obtain knowledge have had a profound influence in the thesis.

Chronologically, one of the first projects and experiences to be highlighted is the work done with the USALPHARMA teaching innovation group. Since 2013 the author collaborates with this group, focused on the application of Virtual Worlds to enrich the learning experiences in the field of Education in Pharmacy and Healthcare. During this collaboration, the author researched about users' interaction and developed a set of tools to be used in the teaching tasks and to be used to analyze the learning and virtual activities performed by the students within the Virtual Worlds environments used.

The second experience chronologically and probably the most important for the thesis was to join the GRIAL Research Group at the University of Salamanca. In this group the author participated in many national and international R&D projects, as well as he had the opportunity of being involved in many collaborations with external partners, participated in some of the most relevant conferences and published in some high-respected journals. During the time spent within GRIAL, the author of this thesis continued participating in the USALPHARMA project. Also, he collaborated with researchers from the Technical University of Madrid in some exciting experiments related to MOOCs (Massive Open Online Courses) as well as he became the principal engineer involved in the development of the Spanish Observatory for University Employability and Employment (OEEU in its Spanish acronym) where he developed some promising research included in this thesis.

While working on the thesis, and related to the grant received by the author from the Junta de Castilla y León in collaboration with the European Social Fund to develop it, the author was involved with the Department of Computer Science and Automation at the University of Salamanca. This involvement was twofold: as a student enrolled in the Ph.D. Programme of Computer Engineering steered by the Department and as a teaching assistant in the subject of Human-Computer Interaction —part of the curriculum in the Computer Engineering degree at that University—.

Finally, and as part of the international research stay required to obtain the International Mention for the Ph.D., the author was a visiting researcher at the IBM Research's T.J. Watson Research Center. The research stay was performed within the IBM Research AI & Q division and was related to software engineering, HCI, and artificial intelligence. During this research stay, the author developed the last case study presented in this thesis.

1.5 Document structure

This document is organized in 6 chapters and 14 appendixes. This section presents the introduction to the whole manuscript, and some of the foundations, motivation and the introduction to the workflow and methodologies followed to develop the research. Chapter 2 presents the theoretical foundation of this research. In particular, a systematic literature review and mapping to contextualize the different experiments carried out in the thesis was developed. Chapter 1 outlines the different experiments performed during the research. Due to the heterogeneity of these experiences and experiments, each one of them is presented separately including the context, material and methods, results and discussion. Chapter 4 discusses the overall findings achieved in this thesis through the different experiments performed, as well as it responds to the research questions posed in the introductory chapter. Chapter 5 presents the overall conclusions of the thesis, some of the future lines of work and the different scientific publications, grants, awards, etc., achieved by the author while carrying out the thesis.

In the case of the appendixes, they include all the papers published with results from the thesis which are relevant in each case study presented in the document.

2 State-of-the-art in systems/software architectures supporting Human-Computer Interaction

2.1 Introduction

Following all the ideas presented in the introduction, the subject of this section —a systematic literature review—, is the combination of software architectures and HCI/HMI processes: how can software architectures support HCI/HMI processes? How could software architectures help, improve, analyze, intervene or contribute to HCI/HMI processes? In the case of this thesis, the literature review addresses these questions from the published literature in these research fields. To improve the literature review, the HMI subfield of HCI has been included to expand the scope of the findings.

To accomplish these goals, and as presented below, was conducted a systematic literature review and mapping following the guidelines proposed by Kitchenham and other authors [114-117].

2.2 Methodology

Kitchenham and other authors [115] have emphasized that an evidence-based approach to software engineering is an important research issue. Finding evidence by reviewing the literature on a specific topic could be addressed using methodologies such as systematic literature reviews and mapping studies [118]. In this context, the primary purpose of a systematic review is to identify, evaluate and interpret the available studies in the literature that take into account the research questions proposed by the authors. Through a systematic review, it is possible to gather evidence to identify gaps and research opportunities in the area of interest. Moreover, the systematic mapping is a form of a systematic literature review that aims to provide an overview by identifying and categorizing the available research on a broad topic based on the guidelines proposed by Kitchenham [115]. Thus, this review is organized using the main activities proposed by Kitchenham: planning, conducting and reporting the study.

2.2.1 Review and mapping planning

When designing the processes of review and planning, the author identified the different objectives, the protocol to follow and other relevant details. In the next sections, an explanation of the different relevant aspects that must be defined for the mapping and review activities to be successful is set forth.

Research questions

The systematic mapping aims to answer the mapping questions (MQ) below:

- MQ1. How many studies were published over the years?
- MQ2. Who are the most active authors in the area?
- MQ3. Which publication vehicles are the main targets for research production in the area? (Journal, conferences, etc.)
- MQ4. In which domains has it been applied? (e.g. Computer Sciences, Education, Medicine, Business)
- MQ5. Which kinds of devices have been involved in software architectures supporting or analyzing HCI / HMI processes (e.g. computers, wearables, smartphones, cameras, etc.)?
- MQ6. What kind of support has been provided by the software architectures to the HMI / HCI processes?

The systematic review aims to answer the following research questions (RQ):

1. RQ1: What are the trends in software architectures that support or analyze Human-Computer Interaction?
2. RQ2: What are the trends in software architectures that support or analyze Human-Machine Interaction?
3. RQ3. Are there significant differences in the trends of software architectures that support or analyze HCI to those applied to HMI?
4. RQ4: What kind of software architectures have been proposed to support or analyze Human-Computer Interaction?
5. RQ5: What kind of software architectures have been proposed to support or analyze Human-Machine Interaction?
6. RQ6: Are there significant differences in the software architectures proposed to support or analyze HCI to those applied to HMI?

Based on the research questions, the author used the PICOC method proposed by Petticrew and Roberts [119] to define the review scope:

7. Population (P): The target group for the investigation. In this study: Software architectures
8. Intervention (I): specifies the investigation aspects or issues of interest for the researchers. In this case, those aspects or issues that provide support or analyze HCI / HMI processes.
9. Comparison (C): the aspect of the investigation with which the intervention is being compared to. No comparison intervention has been planned in the study.
10. Outcomes (O): the effect of the intervention. The goal is to seek for Software Architectures proposals and real-world experiences.

11. Context(C): the setting or environment of the investigation. In this case, they are those environments related to HCI / HMI (in the industry, academia, etc.).

Inclusion and exclusion criteria

To answer the different questions posed, the following inclusion and exclusion criteria were used. The different criteria used to include or exclude a paper were organized into five inclusion criteria (IC) and five exclusion criteria (EC):

- IC1: The papers had a software architecture-based solution AND
- IC2: The presented solution was applied to HCI OR HMI fields AND
- IC3: The presented solution supported OR analyzed HCI OR HMI processes AND
- IC4: The papers were written in English AND
- IC5: The papers were published in peer-reviewed Journals, Books, Conferences or Workshops

The following exclusion criteria were established:

- EC1: The papers did not have a software architecture-based solution OR
- EC2: The presented solution was not applied to HCI OR HMI fields OR
- EC3: The presented solution did not support OR analyzed HCI OR HMI processes OR
- EC4: The papers were not written in English OR
- EC5: The papers were not published in peer-reviewed Journals, Books, Conferences or Workshops

After defining the research and mapping questions (RQs, MQs) and the inclusion and exclusion criteria (ICs, ECs), were defined the sources of the papers and the search string(s) to be used on these sources. The databases were chosen accordingly to the following requirements:

- The database is capable of using logical expressions or a similar mechanism.
- The database allows full-length searches or searches only in specific fields of the works.
- The database is available for the author (through his institution, through his memberships to associations such as IEEE or ACM, which are responsible for some of the databases used, etc.).
- The database is one of the most relevant in the research area of interest in this mapping process: computer science.

Sources: The search was conducted in the following electronic databases:

- Web of Science
- Scopus
- IEEE Xplore

- ACM Digital Library
- Springer Link

Query string

To create the search string, the author identified the principal terms from the research questions, the PICOC and the possible alternative spellings and synonyms. Based on the identified terms, was defined a standard query string using Boolean AND/ OR operators and the (*) wildcard to search for a word with the different possible terminations (plural, singular, etc.). The query was:

("software architectur" AND ((HCI OR "Human-Computer Interaction" OR "Human Computer Interaction") OR (HMI OR "Human-Machine Interaction" OR "Human Machine Interaction")))*

In addition to this query, was defined another for the SCOPUS database, using the proper notation required by this database (thus the logical expression of this different query is basically the same). The query designed for SCOPUS was:

(TITLE-ABS-KEY ("software architectur") AND TITLE-ABS-KEY ((HCI OR "Human-Computer Interaction" OR "Human Computer Interaction") OR (HMI OR "Human-Machine Interaction" OR "Human Machine Interaction")))*

Regarding this query string, the search was not restricted by publication date (the search was conducted including all the papers in time), or other filters provided by the databases. The only unique filter applied was in the Springer Link database, where were avoided all "Preview only" papers, as these papers were not accessible for the researcher using their current subscription. The details of the results obtained using these queries are presented in the following subsection.

2.2.2 Review and mapping process

As previously presented, the search was conducted using the query string described in the previous subsection and with no time (year) restriction regarding the publication date of the papers found. After the search, the selection of papers to be used for the mapping and literature review was performed following these steps:

1. All the results were collected in a GIT repository [120] (in the same manner as they were downloaded from the databases) and in a spreadsheet (https://goo.gl/QK5Qrd), removing all the duplicates across the databases
2. The resultant papers were analyzed based on the title and abstract and the inclusion/exclusion criteria. In those cases where the title and abstract were not sufficient to make a decision, the researcher quickly assessed the entire content of the

paper. The resultant candidate papers were added to another sheet of the spreadsheet document (https://goo.gl/4xweXc)

3. The papers were read in detail and analyzed following the previously-posed research questions. The selected papers passed a quality assessment checklist (see Table 2), and the information was collected in another spreadsheet (https://goo.gl/cYBfyp). Besides the papers selected in this manner, the author also considered papers collected within the references of those that potentially could be interesting for the review and mapping process. This collection of further papers analyzing the references provided the researcher with another 3 papers to review.

Following these steps, the following results were obtained (Figure 2):

1. Execution of the query string in the databases: 2711 papers retrieved (217 from Web of Science, 395 from SCOPUS, 56 from IEEE Xplore, 403 from ACM Digital Library, 1640 from Springer Link).
2. Remove duplicate studies: 2075 (including those that the author did not remove due to the appearance of doubts regarding duplicity).
3. 3.a) Selected papers after reviewing titles and abstracts: 243 (11.71% of the unique papers retrieved). 3.b) Inclusion and review of papers after reviewing the primary paper references: 3
4. Papers selected after reading the full text: 39 (1.88% of the total papers considered, 15.85% of the papers read).

As shown in the guidelines proposed by Kitchenham [115], author formulated a quality checklist to assess the individual studies and avoid subjectivity. These checklists are useful to assist in the papers selection process. The customized quality assessment checklist developed is based on the checklist suggested in [115]. Other works on systematic reviews and mapping of the literature [121, 122] have also customized their quality checklists based on the suggestions given in [115].

Table 2. Quality assessment checklist.

Question	Score
1. Are the research aims related to software architectures & HCI/HMI clearly specified?	Y/N/partial
2. Was the study designed to achieve these aims?	Y/N/partial
3. Are data presented on the evaluation of the proposed solution?	Y/N/partial
4. Are data presented on the assessment regarding the human part of HCI/HMI?	Y/N/partial
5. Is the software architecture clearly described and is its design justified?	Y/N/partial
6. Are the devices involved clearly specified? Are their functions within the software architecture justified?	Y/N/partial
7. Do the researchers discuss any problems with the software architecture described?	Y/N/partial
8. Is the solution based on a software architecture tested in a real context?	Y/N/partial
9. Are the links between data, interpretation, and conclusions made clear?	Y/N/partial
10. Are all research questions answered adequately?	Y/N/partial

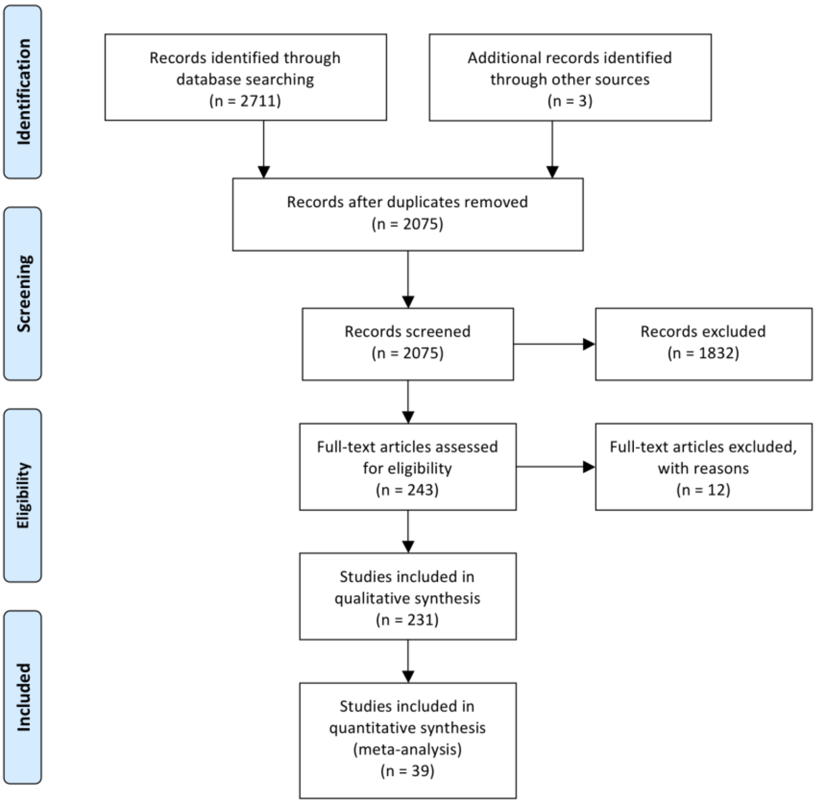


Figure 2. Steps and results of review and mapping process. Reported as proposed in the PRISMA Statement [123].

In the third step of the review, as described above, the papers were read in full, and their quality was evaluated using the quality assessment checklist formulated (Table 2). The answer for each one of the 10 questions could be scored with 1 point if the answer was "Yes", 0.5 points if the answer was "Partial" or 0 if the answer was "No". Using this system, each paper could obtain a score from 0 to 10 points. The first quartile mark (Q1 = 7.5 points or more out of 10 possible) was used as the cutoff score for a paper to be included. If a paper scored less than 7.5, it is excluded from the final list to avoid low-quality works according to the assessment checklist.

2.3 Systematic Literature Mapping results

To map the contents and other aspects contained in the 39 selected papers and to answer the mapping questions (MQ), the papers obtained from the systematic process were analyzed.

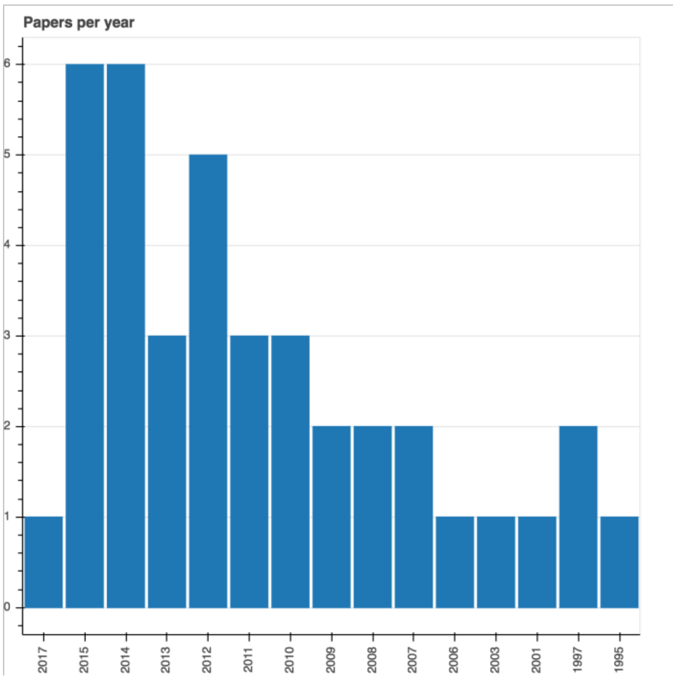


Figure 3. MQ1— Number of papers per year.

Table 3. Authors' names and number of publications

Name	Total
Bellik, Yacine; Bourda, Yolaine; Jacquet, Christophe	2
Abran, Alain; Alenazy, Wael; Almeida, AT.; Alonzo, Rommel; Arkin, C.; Barreto, Guilherme A.; Bass, Len; Behnen, Daniel; Bethel, Cindy L.; Bettini, Claudio; Biel, Bettina; Bigdelou, A.; Bisio, M.; Bongartz, Sara; Bosch, J.; Bozzon, Alessandro; Brambilla, Marco; Brecher, Christian; Bumberger, F.; Burger, B.; Büscher, Christian; Calandra, D.M.; Canedo, L.; Capilla, R.; Cardeñoso, Valentín; Caruso, M.; Carvajal, L.; Caso, A.; Catarci, T.; Chaczko, Zenon; Chan, Cheuk Yan; Chiappalone, M.; Choi, J.; Cincotti, F.; Colajanni, Michele; Coutaz, Joelle; Cremer, Sven; Cutugno, F.; de Almeida Neris, Vânia Paula; de Oliveira Neris, Luciano; Dimakis, Nikolaos; Dorohonceanu, B.; Escudero, David; Ewert, Daniel; Ferrané, I.; Filho, José Tarcísio C.; Folmer, E.; Fraternali, Piero; Garcia Garcia, Daniel; Gill, Raj; Gkekas, Georgios; Gonia, Phillip T.; Gowda, Sandesh; Grill, Thomas; Gruhn, Volker; Habieb-Mammar, Halima; Harrer, Andreas; Hauck, Eckart; Heigemeyr, Andreas; Herfs, Werner; Honorio, N.; HyunRyong Jung; Infantes, G.; Jalaliniya, Shahram; Jarvis, Paul; Jeschke, Sabina; Jin, Yucheng; John, Bonnie E.; Karachristos, Theofilos; Karame, Ghassan; Kates, Jesse; Kausch, Bernhard; Kempf, Tobias; Kim, M.; Kim, S.; Kuz, Sinem; König, Werner A.; Lee, Jamee K.; Leotta, F.; Lerasle, F.; Lin, H.; Lombay-Gonzalez, Jose A.; Lopes, C; Mackin, Michael A.; Malandrino, Delfina; Mardanbegi, Diako; Marsic, I.; Martin-Ruiz, Maria Luisa; Martinoia, S.; Mastromoro, Larry; Mayer, Marcel; Mazzoni, Francesca; Mecella, M.; Mirza, Fahad; Mohamed, Taleb; Moreira, Pedro Miguel; Moshkina, Lilia; Murphy, Robin R.; Müller, Simon; Navab, N.; Nigay, Laurence; Nunes, U; Odenthal, Barbara; Olmedo, Hector; Origlia, A.; ParkRonald, Sunghyun; Parsons, Bernard; Paternò, Fabio; Pau, Ivan; Pires, G.; Pittarello, F.; Polymenakos, Lazaros; Popa, Dan O.; Rego, Paula Alexandra; Reis, Luís Paulo; Reiterer, Harald; Rett, Joerg; Riboni, Daniele; Riccio, A.; Rigoll, G.; Rodrigues Machado, Luciano; Rossi, S.; Rädle, Roman; Santoro, Carmen; Scarano, Vittorio; Schettini, F.; Schilberg, Daniel; Schlick, Christopher M.; Schwarz, L.A.; Seffah, Ahmed; Seoane, Fernando; Silva de Alencar, Tatiana; Simione, L.; Sintos, Ioannis; Soldatos, John K.; Spano, Lucio Davide; Speroni, Paolo; Srinivasan, Vasant; Sutcliffe, Alistair; Tessadori, J.; Thew, Sarah; Tiefenbacher, P.; Toffetti, Giovanni; Tsolakidis, Stamatis; Vega-Barbas, Mario; Veloso, Marcus V. D.; Wang, L.; Warner, Peter; White, Anthony.	1

Figure 3 answers the mapping question MQ1 (*how many studies were published over the years?*) in a graphical way. Despite the systematic mapping process not restricting any range of years, the resultant papers were published from 1995 to 2017. This could be because between 1980 and 1990, some of the foundational papers and books on Software Architectures [15, 16, 124] and Human-Computer Interaction [125, 126] were published, so the intersection between the two fields appeared in later years.

To answer MQ2 (*Who are the most active authors in the area?*), the authors from the selected papers were identified. Only three researchers appeared more than one time in the results (they are co-authors in two of the resultant papers). Thus, the 39 papers included 151 different authors. The full list of authors is shown in Table 3.

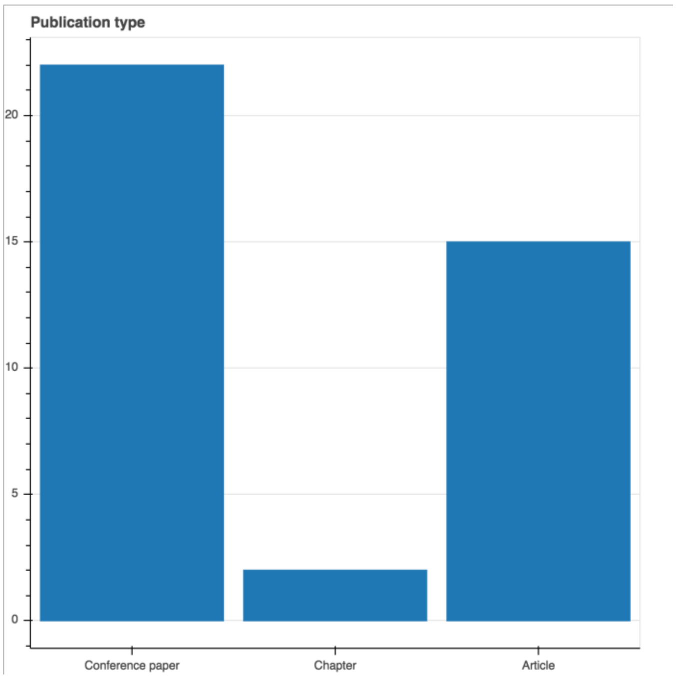


Figure 4. MQ3— Publication type (channel).

Regarding MQ3 (*which publication vehicles are the main targets for research production in the area? Journal, conferences, etc.*), Figure 4 shows the publication channel of the selected papers. Most papers, 22/39, were published at conferences [17, 127-147] (56.41%), 15 papers [148-162] (38.46%) were published in journals, and the remaining 2 papers were published as book chapters [163, 164] (5.13%). The higher number of papers published in conferences could be explained as other authors note [121], because "in Computer Science, the number of conferences is significantly higher compared with the number of journals, which explains why most papers were published through this channel". To complete the MQ3, the name of the conference, book or journal that published the selected papers is provided in Table 4 along with their h-index according to the SCIMAGO index (SJR) and the number of selected papers published in each of them. This could help with the understanding of the primary sources regarding software architecture & HCI or HMI as well as their impact on the scientific community.

Table 4. Publication source

Reference	Source name	*h-index*
[135]	ACM International Conference on Intelligent User Interfaces	50
[133]	ACM International Conference on Multimodal Interaction	13
[142]	ACM International Joint Conference on Pervasive and Ubiquitous Computing and ACM International Symposium on Wearable Computers	9
[163]	Advanced Intelligent Environments	-
[147]	Ambient Intelligence	-
[144]	Automation, Communication and Cybernetics in Science and Engineering 2011/2012	-
[157]	Autonomous Robots	88
[140]	Conference on Human Factors in Computing Systems	131
[152]	Frontiers in Neural Circuits	36
[136]	Human-Computer Interaction – INTERACT	-
[127]	IBIMA Conference	3
[165]	IEEE Aerospace Conference	41
[143]	IEEE Emerging Technology and Factory Automation (ETFA)	
[131]	IEEE International Symposium on Industrial Electronics	42
[137]	IET International Conference on Intelligent Environments	10
[150]	International Journal Of Ad Hoc And Ubiquitous Computing	16
[130]	International Conference on Advanced Networking and Applications	36
[139]	International Conference on Advanced Robotics	16
[141]	International Conference on Automotive User Interfaces and Interactive Vehicular Applications	5
[132, 134]	International Conference on Human-Computer Interaction, HCII	-
[17]	International Conference on Software Engineering	105
[154]	International Journal of Human-Computer Interaction	51
[153]	International Journal of Information Technology and Web Engineering	7
[158, 160]	International Journal of Social Robotics	29
[149, 151]	Journal of Systems and Software	82
[161]	Journal of Intelligent & Robotic Systems	50
[155, 156]	Journal on Multimodal User Interfaces	17
[146]	New Perspectives in Information Systems and Technologies	-
[128]	Next-Generation Robotics II; and Machine Intelligence and Bio-inspired Computation: Theory and Applications IX	-
[159]	Personal and Ubiquitous Computing	68
[164]	Relating System Quality and Software Architecture	
[162]	Requirements Engineering	43
[148]	SENSORS	104
[138]	SICE-ICASE International Joint Conference	19
[145]	Universal Access in Human-Computer Interaction. Aging and Assistive Environments	-

In the case of question MQ4 (*in which domains has it been applied? E.g., Computer Sciences, Education, Medicine, Business),* the selected papers proposed software architectures to solve HCI/HMI issues related to different application domains. As illustrated in Figure 5, the software architectures were applied to domains such as aeronautics [129] (2.56%), BCI systems [134, 152] (5.12%), cognitive systems [144] (2.56%), collaboration systems [154] (2.56%), context-aware systems [136, 150, 159] (7.7%), eyewear computing [142] (2.56%), healthcare systems [131, 135, 146, 148] (10.25%), industrial machines [130, 132] (5.12%), mobile systems [143, 149, 164] (7.7%), multimodal systems [128, 133, 137, 140, 155, 156, 163] (17.95%), robotics [138, 139, 157, 158, 160, 161] (15.38%), smart environments [147] (2.56%), software engineering [17, 162] (8.16%), ubiquitous environments [145] (2.56%), vehicle systems [141] (2.56%) or eLearning systems [127] (2.56%).

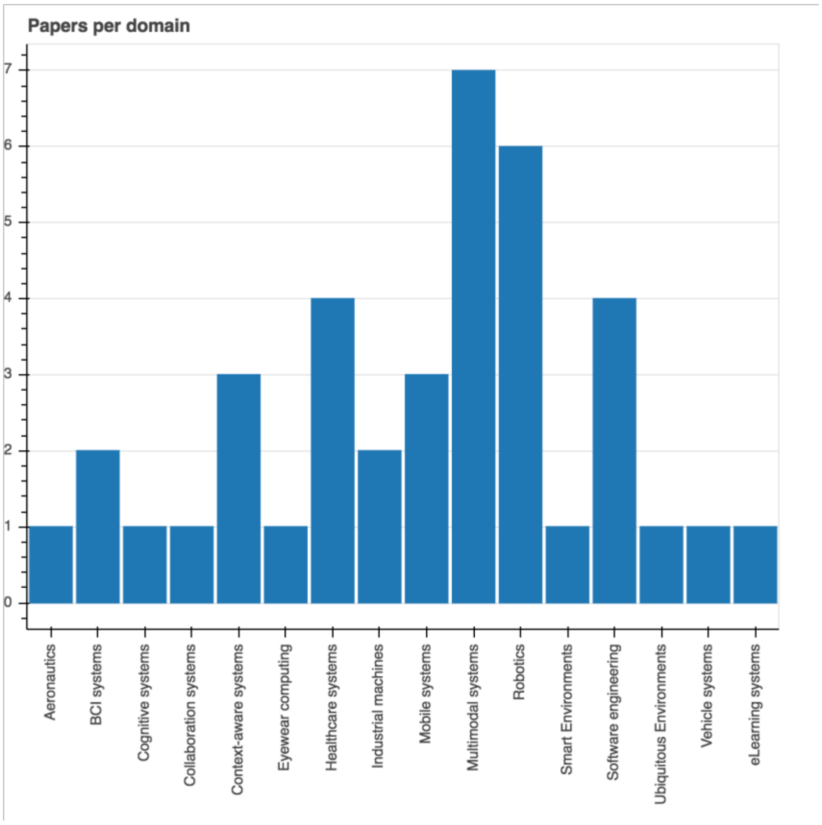


Figure 5. MQ4— Application domain.

As for question MQ5 (*which kinds of devices have been involved in software architectures supporting or analyzing HCI / HMI processes? E.g., computers, wearables, smartphones, cameras, etc.),* the selected papers showed different devices that were supported or included in the software architectures proposed, with those shown in Table 5. In this case, the mobile devices, the robots, or other input sources such as a camera, microphones, gaze trackers, etc., stood out. These results are directly related to the results from question MQ6 because, as shown below, the software architectures presented in the papers mainly supported multimodal or human-robot interactions. Readers can check https://goo.gl/yEq6TG to get further information about what paper (and software architecture proposal) included each device.

Table 5. Devices used in the software architectures proposed.

Name	Total
Mobile devices & smartphones	12
Robots	9
Camera	6
Computers	6
Microphone	4
Gaze tracker	3
Industrial machines	3
Sensors	3
Speech input systems	3
Brain-Computer Interfaces (BCIs)	2
Tablet	2
Touch screen	2
Wearable	2
Avionics box	1
Biosensors&BCI	1
Body tracker	1
Cuff display	1
Custom-build table	1
Digital Home	1
Display & Control module	1
Dynamic displays	1
Eyewear computers	1
Force pressure sensors	1
Force-Feedback tactile gloves	1
Google Glass	1
Home appliances	1
LEAP	1
Maps viewer	1
Microsoft surface	1
Motion tracking sensors	1
Mouse, keyboard & Joystick	1
Raspberry Pi	1
Server	1
Smart board	1
Smart projector	1
Smartphone	1
Speaker	1
Touch sensors	1
Vehicles	1
Wheelchair	1

Regarding question MQ6 (*what kind of support has been provided by the software architectures to the HMI / HCI processes?*), the papers showed different kinds of support provided by the software architectures. As illustrated in Figure 6, the software architectures included support to varying HCI/HMI processes, including the following: 7 papers [127, 128, 135, 140, 147, 155, 157] were related to the multimodal interaction support provided by software architectures (17.95% of the papers), in 6 papers [138, 139, 144, 158, 160, 161] the software architectures supported human-robot interaction (15.38%), in 5 papers [17, 149, 151, 153, 164] these aimed to improve usability (12.82%), in another 5 [130, 131, 134, 141, 152] they supported human-machine interaction (12.82%), in 4 papers [136, 145, 148, 159] software architectures were related to articulate context-awareness (10.26%), in 3 [137, 150, 163] they defined software agents to adapt software architectures to different contexts (7.69%), in 1 paper [129] the software architecture supported the proper connection of software and hardware components (2.56%), in [146] it defined interaction models to use with 3D contents (2.56%), in [132] it supported the evaluation of touch interaction (2.56%), in [142] it supported eye interaction with a system (2.56%), in [162] it aimed to support the integration of software requirements and HCI (2.56%), [133] aimed to analyze multimodal interactions (2.56%), in [154] the software architecture supported multimodal interaction for user collaboration (2.56%) and lastly, in [156] the software architecture supported multimodal interaction with 3D environments (2.56%). In the case of this question, the categories regarding the support provided have been defined using the leading software architecture's purpose outlined in the paper.

Table 6 summarizes the main findings of the mapping report. These findings could only be considered within the scope of this research and its conditions. They do not represent the whole state-of-the-art in the mapping of the intersection between software architectures and HCI research fields.

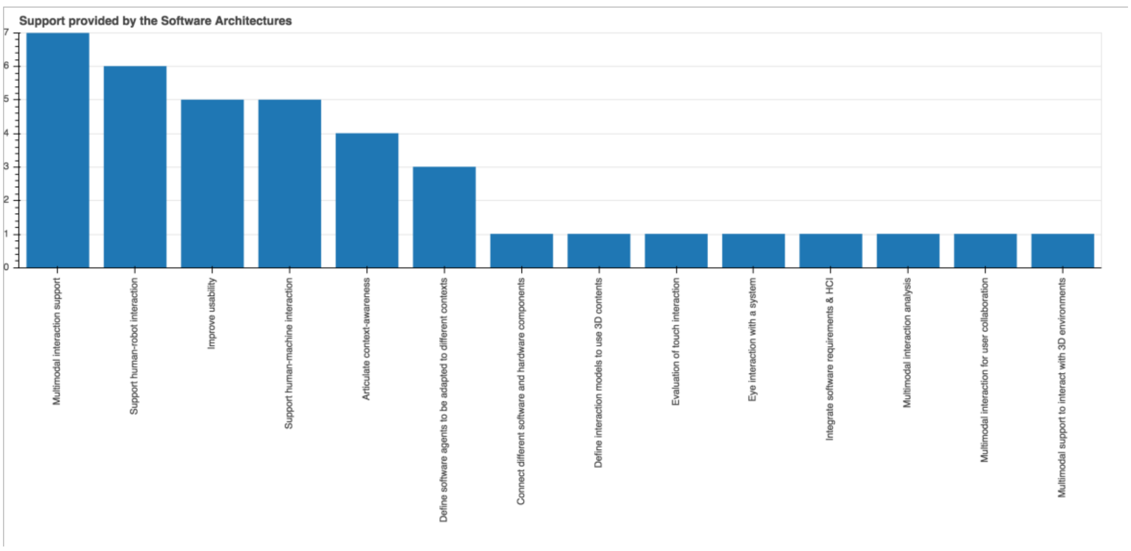


Figure 6. MQ6— Support provided by the software architectures.

Table 6. Summary of the mapping report findings.

Mapping question	Finding
MQ1	The papers were published between 1995 - 2017
MQ2	The most active researchers are Yacine Bellik, Yolaine Bourda and Christophe Jacquet
MQ3	Most papers are published in conferences
MQ4	The main application domain is in multimodal systems
MQ5	The principal devices involved are the mobile devices and smartphones, followed by the robots (and robotic systems)
MQ6	The main purpose of the software architectures presented in the resultant papers is to support multimodal interaction

2.4 Systematic Literature Review Results

To begin this section, would be good to remark on the difference explained in the first section (introduction) of this chapter between HCI and HMI. In this case, the HMI was considered as a subset within the HCI field. HCI can be considered as the main field, which takes into account all kinds of interaction between the human and the interactive computing systems; while HMI is the internal field where only the interactions between the human and the machines that are not traditional computers (but probably contain them) are considered. In this case, a clear example of such difference are robots, which are considered within the HMI scenario. Robots, despite their reliance on complex computing systems, have the shape of a machine and its classical attributions and peculiarities, in contrast to other kinds of computers such as PCs, smartphones, tablets, wearables, etc. The difference between HCI/HMI in this thesis is not based only on whether the artifacts or devices within which the software architecture acts upon have or not a graphical user interface (GUI). The difference is based on whether the software architecture supports the interaction in elements or devices that could be framed as a computer or as a machine (used for aims other than as a sole computer). That difference is crucial for answering the following questions. All robotic systems and industrial machines are considered in the subfield of HMI, and all the other computing systems are considered in the field of HCI despite them using input/output devices that are mere machines. As a result, the papers [17, 127, 129, 132-134, 136, 137, 140, 142, 143, 145-151, 153-156, 159, 162-164] are framed within the field of HCI while the other [128, 130, 131, 135, 138, 139, 141, 144, 152, 157, 158, 160, 161] are considered as part of the field of HMI in the following questions.

Regarding RQ1 (*what are the trends in software architectures that support or analyze Human-Computer Interaction?*), first of all, it should be noted that the software architectures presented are intended mostly to support interaction, not to analyze it. The only paper that clearly presented a software architecture that was fully intended to support interaction analysis was [133]. That paper presented a software architecture that supported and analyzed interaction with a multimodal system. In other cases, as commented below, the papers analyzed user interaction as part of their features to decrypt users' inputs (but the analysis was not their main goal). Moreover, Table 8 and Table 9 present the main interfaces used in the systems where the software architectures work and the functions related to Human-Computer Interaction that those software architectures support.

In the case of the user interfaces (Table 8), most systems and architectures used GUIs (Graphical User Interface) as the contact point with the user 15/26 (57.69% of the papers), followed by mobile devices, smartphones or tablets that were the main physical interface presented (in 12/26 papers, 46.15%). The other interfaces (to receive the users' input) included in the resultant papers were are microphones and other speech input interfaces (in 30.77% of the papers), cameras (26.92%), motion tracking sensors (19.23%), touch display or touchable interfaces (19.23%), mouse, keyboard or joysticks (15.38%), gaze trackers (11.54%), wearables (11.54%), Brain-Computer interfaces (BCIs, 7.69%), force pressure sensors and haptic devices (7.69%), "laser pointers" (in 3.85% of the papers).

As commented, Table 9 shows the primary findings on the logical support provided by the software architectures. The main goal for the software architectures in the selected papers was related to improving usability (present in 34.62% of the papers), followed by the support of multimodal interaction (26.92%) and information fusion from multiple sources (also in 26.92% of the papers). In the other papers, the aim of the software architectures to support HCI processes were to support the usage of non-classical interaction devices or systems (11.54%), to support context-awareness features or interaction (11.54%), to support adaptability to the user (11.54%), to enhance the collaboration between users (3.85%), to evaluate user interaction or gestures (3.85%), to analyze user behavior or profile (3.85%), to enhance accessibility (3.85%), to help in providing assistive services (3.85%) or to communicate software and hardware systems between users and external operators (also in 3.85% of the papers).

Under the scope of trends, the topic of the interfaces used was clearly focused on mobile technology and GUIs (Table 8). Most papers from 2009 onwards (11/19 -57.89%- papers published in those years) were related to smartphones, mobile devices, and tablets as input sources for user interaction. Furthermore, GUIs were included in 57.69% of publications across all years (1995-2015). In spite of this, in recent years (2014 and 2015) fewer papers were found dealing with GUIs (5/9, 55.56%, of papers published in 2014-2015) as other terms and interfaces became interesting for the researchers such as the so-called *natural interfaces*. The other input sources and interfaces were more balanced across the entire time span of publications. Other short-trends appeared, such as the case of two papers [134, 146] that dealt with BCIs, published in 2013 and 2014, respectively.

In the case of the support provided by the software architectures, there were no clear trends in general. The only specific trend that could be observed was that all the software architectures that were focused on supporting the interaction with non-classical devices or systems were published in 2015 [127, 142, 156]. Related to other kinds of support provided by the architectures in the literature, as observed in Table 9, either there are few papers that used a specific kind of support or the topics did not follow a trend (i.e. some kinds of support appeared every year included in the resultant papers, so they did not follow a trend). As an example, the papers that investigated software architectures that supported the improvement of usability were recurrent over the years (this aim appeared in 10/21 papers published between 2001 and 2014). Also, software architectures that supported the fusion of information from different inputs were a frequent subject in papers from 1995 to 2015 (there were no periods longer than 3 years without a paper on this topic).

As a final comment related to trends, and as commented in the following RQ4 and Table 11, in the case of HCI papers it was clear that the layered architectures were popular in recent years (2014-2015), with 4 out of 9 papers [127, 132, 143, 164] including it. Also, there was a clear trend in the usage of software architectures based on agents. All papers related to this approach were published before 2010 [137, 140, 150, 154, 163]. Lastly, the last trend to be highlighted is the one related to architectures based on web services: 5/7 papers that covered this topic were published between 2010 and 2012 [129, 136, 147, 149, 159].

Table 7. Summary of review findings on what are the main interfaces/systems presented in HMI papers.

Reference	Robotic system	Multimodal interface (gestures and body recognition input)	Multimodal interface (speech input)	Multimodal interface (pressure sensors, haptics, etc.)	Other sensors (humidity, temperature, etc.)	Graphical User Interface	Brain-Computer Interface (BCI)
[128]	X	X		X		X	
[130]						X	
[131]			X			X	
[152]						X	X
[135]		X				X	
[138]	X					X	
[139]	X				X		
[141]							
[144]		X				X	
[157]	X	X	X				
[158]	X	X	X		X		
[160]	X	X	X		X		
[161]	X	X				X	

Table 8. Summary of review findings on the main interfaces/systems presented in HCI papers.

Reference	Graphical User Interface	Laser Pointer	Touch display / interface	Camera	Microphone	BCIs (Brain-Computer interface)	Motion tracking sensors	Mouse, Keyboard & Joystick	Force pressure sensors & Haptics	Smartphone /mobile devices/ tablet	Gaze trackers	Wearables
[148]												
[127]				X			X		X			
[129]			X	X								
[149]										X		
[163]	X		X	X	X		X			X		
[150]	X			X	X		X					
[151]	X											
[164]	X									X		
[132]	X		X									
[133]	X		X		X						X	
[134]	X		X			X				X		
[136]	X									X		
[153]												
[137]				X								
[154]	X			X	X			X		X	X	X
[140]	X							X				
[17]												
[142]										X	X	X
[143]	X									X		
[155]	X	X			X					X		
[156]	X				X			X				
[145]										X		
[146]				X	X	X	X	X	X			
[159]										X		
[147]	X				X		X			X		X
[162]	X											

Table 9. Summary of review findings on the logical support provided by the software architecture to the interaction process in HCI papers.

Reference	Support non-classical interaction devices or systems	Context-aware	Fusion input information from multiple sources	Multimodal interaction support	Improve usability	Adaptability	Enhance user collaboration	Evaluate interaction / gestures	Analyze user behavior or profile	Accessibility	Assistive services	Communicate software and hardware systems between user and external operators
[148]			X								X	
[127]	X		X	X								
[129]												X
[149]					X							
[163]			X	X					X			
[150]		X		X								
[151]					X							
[164]					X							
[132]								X				
[133]				X								
[134]					X					X		
[136]		X										
[153]					X							
[137]			X	X								
[154]			X		X	X	X					
[140]			X	X								
[17]					X							
[142]	X											
[143]					X							
[155]												
[156]	X											
[145]						X						
[146]			X	X								
[159]		X										
[147]						X						
[162]					X							

Table 10. Summary of review findings on the logical support provided by the software architecture to the interaction process in HMI.

Reference	Interaction analysis using Artificial Intelligence techniques	Interaction analysis using statistical techniques	Feelings, behavior, interest recognition	Information fusion	Personalization	Multimodal interaction recognition
[128]						X
[130]					X	
[131]			X			
[152]						
[135]		X		X		X
[138]			X			
[139]	X					
[141]				X		
[144]						X
[157]		X		X		X
[158]			X			X
[160]			X			X
[161]						X

In the case of RQ2 (what are the trends in software architectures that support or analyze Human-Machine Interaction?), Table 7, and Table 10 show the different interfaces or interaction systems found in the HMI-related papers, as well as the support provided by each software architecture outlined. In general, most of the papers referred to robotic systems that use different input sources for human interaction (7/13 papers, 53.85%). In general, most of the architectures were designed to support interaction via two main interfaces: graphical user interfaces (61.54% of the papers) and/or multimodal interfaces intended to recognize users' gestures or body —typically cameras but also using other devices such Kinect, etc.— (53.85% of papers). Besides these main two interfaces employed, other interfaces such as microphones and other speech input devices for multimodal purposes (30.77%), pressure sensors for multimodal interaction (7.7%), brain-computer interfaces (BCIs, 7.7%), or general-purpose sensors such as those for humidity, temperature, etc. (23.08% of papers), appeared in the selected papers.

In the case of the support, the main one provided by the software architectures was related to supporting the recognition of multimodal interaction (7/13 papers, 53.85%). In the other cases, the support provided was related to helping in recognizing users' feelings, behaviors, or interests using mostly ad-hoc techniques (30.77% of papers), helping in the application of information fusion techniques (23.08%), supporting the application of statistical techniques to analyze interaction (15.38%) such as Hidden Markov Models [135, 157], using artificial intelligence (neuronal networks [139]) methods to analyze interaction (7.7%) or supporting personalization in the human-machine interaction (7.7% of the papers).

As for the trends themselves, the following can be said: there was a clear trend in interfaces working with multimodal interfaces and systems. From the point of view of the interfaces involved, most papers from 2011 dealt with multimodal challenges in several ways: 7/13 (53.85% [128, 135, 144, 157, 160, 161]) of total papers –77.78% of the papers published since 2011– used multimodal interfaces such as gesture and body recognition system, speech inputs (for example, only [131] dealt with speech input before 2011) or pressure sensors (1/13 of the papers). Regarding the support that software architectures provided to the HMI processes, the multimodal interaction recognition was also the trending topic from 2011. In this case, the same papers that used multimodal interfaces ([128, 135, 144, 157, 160, 161]) –77.78% of the papers published since 2011; 53.85% of all papers in HMI category–, dealt with a software architecture in charge of recognizing the multimodal interaction. Regarding other trends, it should be noted that software architectures focused on recognizing feelings, behavior or interests between users were trendy between 1997-2011 (3 papers out of the 5 published in that time span discussed this). With respects to other interfaces or support provided by the software architectures presented in Table 7 and Table 10, the remaining contents were not grouped into particularly interesting trends, since most of them were commented in the papers without constituting them as a particular trend.

As a final remark, there was a clear trend in the software architectures structure of HMI papers (commented in RQ5 and Table 12), as most papers published before 2015 (70%) were based on modules or components (mostly in monolithic compositions), and papers from 2007 to 2017 (60%) were based on layered architectures (with similarities in the usage of both structural approaches between 2007-2014).

As related to both previous questions, RQ1 and RQ2, the RQ3 question (*are there significant differences in the trends of software architectures that support or analyze HCI to those applied to HMI?*) are answered.

In this section, is briefed a comparison between the different main aspects analyzed in the previous RQ1 and RQ2 questions.

Concerning the interfaces mentioned in each paper that studied HCI and HMI, it was observed that the Graphical User Interfaces were highly relevant, as more than 50% of the papers in each area covered GUIs as the main interaction interface for the architectures presented. The main difference in their trends was that in the case of HCI-related papers, the main interfaces were related to mobile devices (smartphones, tablets, etc.), while in HMI the main interfaces were those related to recognizing users' gestures or body position. In the case of HCI papers, the interfaces that collected users' gestures also appeared (under the motion tracking and camera terms). In the case of HMI papers, mobile devices did not appear *per se.* In HMI papers the term GUI appeared, but not in mobile devices, although it was present in desktop computers, in embedded systems, etc. Among the less *trendy* interfaces, similarities such as the appearance of BCI systems in both fields were found.

When discussing the support provided by the software architecture to the HCI/HMI processes, as previously explained, in the field of HCI-related papers, there were no clear trends (most of the types of supports appeared a few times or were constant over the time). While in the HMI-related papers there was a clear trend: 77.78% of the papers dealt with the recognition of multimodal interaction.

Lastly, in the case of software architectures, even though it is commented in depth below, can be highlighted some common trends in HCI and HMI papers: in recent years, layered architectures became relevant for the researchers in the field of software architectures. In HCI papers they became popular in 2014-2015; in contrast, in HMI they were broadly used since 2007 and are continuously used nowadays. As for the other approaches and structures presented, there were fewer coincidences: in both cases, software architectures based on components and modules or in agents were presented, but with no evident coincidence in the time regarding their popularity. Also, in HMI papers the software architectures based on web services did not appear, while in HCI papers, there was an increasing trend in recent years (2015 mostly).

To answer RQ4 (what kind of software architectures have been proposed to support or analyze Human-Computer Interaction?), Table 11 is shown below, which briefly outlines the main findings on how software architectures in the selected papers were defined. As a result of the review, was found that most of the architecture proposals were based on modules and components –generally isolated– (in 38.46% of the papers). Regarding the composition, the other preferred structures were those based on layers (19.23%) or agents (19.23% also). The proposed architectures, in general, were *ad-hoc* and did not explicitly follow previous proposals. The exceptions were: the paper from Nigay and Coutaz [140] that based their proposal on the PAC-Amodeus [166], the chapter by Jacquet, Bourda and Bellik, [163] that was based on their previous work on the KUP model for multimodal interaction [137] and the work by Bongartz, Jin, Paternò, Rett, Santoro and Spano [147] that used the MARIA model-based framework [167].

Regarding other aspects, there was a clear trend regarding the use of web services to articulate the communication between internal software entities or external systems (present in 26.92% of the papers). In fact, 42.3% of the HCI-related papers reviewed (11/26) clearly presented the communication protocols or data structures employed in the software architectures. Another aspect to highlight is that 5/26 of the papers (19.23%) were theoretical proposals or research studies in the field of software engineering. For that reason, they did not provide architectures *explicitly* but provided handy explanations on how the software architectures affect different aspects related to HCI (most of them discussed how structural aspects could help improve usability, etc.). Finally, the last feature extracted from the reviewed papers was that only one of them (3.85%) used ontologies to define the software architecture behavior or services.

In the case of RQ5 (*what kind of software architectures have been proposed to support or analyze Human-Machine Interaction?*), Table 12 presents the main findings related to the structure and features of the software architectures outlined in the papers. In this case, was found that most of the architectures were based on modules or components (53.85%) and these were mainly under the scope of monolithic architectural designs. Moreover, there were other papers (6/13, 46.15%) that presented layered architectures where the different parts were not framed within modules or simple components. In both approaches (architectures based on components and those layered) sometimes (in 3 architectural proposals, 23.08%) coordinator entities that manage all the architecture and coordinate the interaction between the different components or layers were presented. In one case, the architecture was designed as a layered architecture and also as an architecture ready to be used in a distributed way between different devices and systems (using APIs and other web-based network protocols) [130]. Regarding the information flows and protocols used by the architecture to communicate with the internal software entities or to communicate with external stakeholders or software systems, only 6 out of the 13 (46.15%) proposed architectures specified their communication protocols, data formats, etc. (most of them used XML files to transmit information between components and software entities). However, in the remaining cases, such a fundamental part is obviated. Also in general, the papers did not present traditional architectures (PAC, PAC-Amodeus, MVC, etc.) as the basis for their proposal, but they presented *ad-hoc* approaches.

Lastly, RQ4 and RQ5 were compared to answer RQ6 (*are there significant differences in the software architectures proposed to support or analyze HCI to those applied to HMI?*). This question is related to all previous ones in this section. There were some coincidences in the structures, composition, and design presented in the papers. Both in HCI and HMI-related papers, architectures composed of modules or components, based on agents, layered ones, etc. were presented. Also, in HCI software architectures that established their activity and communicated with each layer, agent or component using web services were presented. In the case of HMI papers, this issue was not widely covered (only one paper mentions it).

Table 11. Summary of review findings on software architecture definition in HCI papers.

Reference	Theoretical proposal / study	Architecture based on modules and components	Architecture based on agents	Layered architecture	Architecture based on web services	Uses ontologies to define software behavior or services	Defines communication protocols or structures
[148]		X			X		X
[127]				X	X		
[129]					X		X
[149]	X				X		
[163]			X				
[150]			X				X
[151]	X						
[164]	X			X			
[132]				X			X
[133]		X					X
[134]		X					
[136]		X			X		
[153]				X			X
[137]			X				
[154]			X				
[140]			X				
[17]	X						
[142]		X					
[143]				X			
[155]		X					X
[156]		X					X
[145]						X	X
[146]		X					
[159]		X			X		X
[147]		X			X		X
[162]	X						

Also, in the case of HCI-related papers, there were some theoretical studies (in the field of software engineering) that tried to standardize different common issues or deal with specific problems from a logical and pattern perspective. In the HMI papers, there were no theoretical approaches, as these papers were more focused on giving practical applications to their context and challenges. Also, the software architectures that appeared in the HCI category had more complex architectures that comprised more components, behaviors, and uses, while in HMI the architectures were simpler and mainly had the purpose of dealing with multimodal challenges mostly from a physical perspective. Following this idea, in general, the papers related to HCI had richer descriptions of their software architectures, communication protocols, etc., than those present in the HMI category.

Table 12. Summary of review findings on software architecture definition in HMI papers.

Reference	Architecture with a coordinator entity	Architecture based on modules or components	Layered architecture	Distributed architecture	Define the communication protocols?
[128]			X		
[130]			X	X	X
[131]		X			
[152]		X			
[135]		X			X
[138]		X			X
[139]		X			
[141]		X			X
[144]	X		X		
[157]			X		
[158]	X	X			
[160]	X		X		X
[161]			X		X

To finish, it could be remarked that most of the architectures presented (at least those that appeared in practical papers) were *ad-hoc* and did not follow well-known models or approaches that were previously present in the literature. Exceptions were three papers in the HCI category that worked with previously-defined architectures such as PAC-Amodeus [140], MARIA framework [147] and with KUP models [163].

As a coda to this section, Table 13 shows a summary of the main findings of the systematic review report.

Table 13. Summary of the systematic review report findings.

Review question	Finding
RQ1	Most HCI papers outlined GUIs as the main interface with the user (57.69% of the papers), followed by mobile devices (46.15%). Regarding trends in interfaces and support aim, it should be noted that most papers from 2009 onwards involved mobile devices in general (57.89% of papers), and an important part of the papers within the 1995-2015 period used GUIs as the main non-physical interface. The main support provided by the different software architectures presented was to improve the usability of the systems. In the case of the support provided, there were no relevant trends over the years, despite some kinds of support appearing more than others in the literature. Regarding the software architectures definition, design, structure, etc., the papers presented layered architectures as a trend in recent years (2014-2015), while the trends in papers published before 2010 were agent-based architectures and in papers published between 2010 and 2012, these were based on web services.
RQ2	Most HMI-related papers also established GUIs as the main interface (61.54%), followed by multimodal interfaces to recognize users' gestures or body position (cameras, Kinect, etc., 53.85%). In this case, the systems managed by the software architectures consisted on machines with robotic shapes (53.85%). Regarding the main aim of the software architectures presented in the papers, the recognition of multimodal interaction stood out. In the case of the architectures' design or structures, the trends were related to: before 2015, most of the papers presented components or module-based architectures (mainly monolithic ones). In recent years (2007-2017) a growing trend on using layered architectures was observed.
RQ3	GUIs had a similar relevance in both paper categories. Also, in HCI-related papers the mobile devices appeared as the second-most relevant interface, while in those related to HMI the second-most relevant interfaces were those related to gestures or body recognition. As for the support provided by the software architectures, in HCI papers there were no relevant trends. In the HMI-related ones, the trend found was the support in recognizing users' multimodal interaction (77.78% of papers). In the case of the software architectures structure and design, in both categories layered architectures appeared as relevant in recent years. Other individual trends, such as the architectures based on modules or agents, exhibited no clear coincidences over time.
RQ4	Most of software architectures proposed were based on modules or components (38.46%). The other main proposals were based on layers (19.23%) or agents (19.23%). Only three software architectures were not purely *ad-hoc* and were based on previous works. Considering other trends in the structure and design, 26.92% of the software architectures based the communication between their components or with external systems on web services.
RQ5	In the case of software architectures design or structure in HMI-related papers, 53.85% were based on modules or components (mainly in monolithic shapes). Also, the other 46.15% of the papers presented layered architectures. In general, in this case, 46.15% of the papers specified their communication protocols, although only one used web services to communicate with components.
RQ6	In both HCI and HMI-related papers, there were architectural designs that were based on components or modules, agents, or layers. Software architectures that base their communication on web services were present in HCI papers (only 1 in the case of HMI). It should be noted that in HCI papers there were some theoretical works on how software architectures could improve usability or other issues, while in the HMI ones this kind of work was not present.

2.5 Analysis of proposed solutions

In this section are briefly commented the different proposals made by the authors across the 39 selected papers. To comment the proposals will be used the different kinds of support provided by the software architectures (as presented in MQ6 – Fig. 5 in a simple classification and RQ1 – Table 9 and RQ2 – Table 10 in a richer way) as the basis for articulating the comments.

In the case of the papers that included software architectures intended to purely support multimodal interaction [127, 128, 135, 140, 147, 155, 157], the authors proposed the following different challenges and approaches. In [127], Chaczko, Alenazy, and Chan presented a way to enhance smart learning environments through the use of software architecture (in this case described as middleware) that controls and manages all the different interfaces and resources available, such as haptic interfaces, cameras, and gesture/body trackers, etc. The middleware they used manages the network communication, integration of all inputs from different devices, etc., to combine all the interactions and gather them into a combined mode to be tracked and used within the smart learning environment software platform. Regarding the evaluation of the proposal, the authors presented some preliminary experiments to report on how the software architecture with the gestures or body recognition as the input, could work. The paper by Alonzo *et al.* [128], also presented a system intended to combine the users' interaction with multimodal sensors, but in this case in the context of personal robotics. Despite its context being related to the use of personal robotics environments, the paper was categorized mainly within the support of multimodal interaction since the authors focused the main findings and proposals on a logical architecture that combines the different inputs. To combine the different inputs, the paper dealt with some custom algorithms and techniques to recognize and manage interaction from different sources, and it presented some evaluation on their experience implementing the proposal as well. The paper by Bigdelou, Schwarz, and Navab [135] also presented interesting work on how to improve the multimodal interaction (gesture-based mostly) in the case of modern clinical and healthcare contexts. In this case, their proposal was not only focused on recognizing the users' gestures or other kinds of interaction, but also on learning from them by applying machine learning techniques. In the case of the software architecture, they presented a system that handled the interaction (and learn from them) to later use the resultant information as processed inputs for healthcare platforms and software. Regarding the tests and practical application of their work, the authors made a valuable effort in the paper to compare different gesture recognition techniques and the users' intentions and perceptions about the experience using the system. In [140], Nigay and Coutaz explained an approach of combining the agent-based PAC-Amodeus model with a fusion mechanism to support a multimodal airline travel information system (named MATIS). In the paper, the authors explained the different software components from the PAC-Amodeus model (a previous model developed by Coutaz and others) and how they are combined with a custom fusion engine and algorithms to respond to the users' interaction with the dialog controllers. As a proof of concept, the authors provided an example of how the interaction could be conducted by the users (using speech inputs and mouse buttons, for example) and how it could manage this interaction to give feedback to the user about the interaction to thereby successfully use the travel information system. The paper by Bongartz *et al.* [147], presented a research work on how to provide adaptive user interfaces to users within the context of smart environments. In the paper, the authors explained how they used the MARIA model-based framework to obtain adaptation capabilities that could be used in various interaction modalities. They used the different languages (to descript interfaces, to define the different devices involved, etc.) provided by MARIA to generate a proof of concept system and different heuristics to be applied in the adaptation to the user in the case of a mobile application and wearable interfaces (a display, microphone, earphone, etc.). To conclude their study, the authors present an evaluation of the interaction and user experience based on a survey of 10 participants. In another work, König, Rädle, and Reiterer [155] presented a pilot experience with different interfaces and devices (again aligned with the multimodal interaction paradigm) to design new interactive multimodal experiences. In this case, they provided interesting details on data types, information combination approaches, visualizations and GUIs to design the system, etc. As a proof of concept, the authors present how they used their approach to build a multimodal interaction system employed in an exhibition including laser pointers, mobile eye tracking, Nintendo Wii devices, 360º panoramic screens, etc. Regarding the evaluation of the system, the authors explained different evaluations to be used and overall results of the proof of concept, but they did not provide specific data about that task. To finish with the papers that were fully focused on supporting multimodal interaction, Burger, Ferrané, Lerasle and Infantes [157] presented a work on how to deal with gesture interaction (in this case using two hands) and speech to control and interact with robots. In this case, the authors employed a software architecture that encapsulates the recognition of the interaction with each device or tracker (using algorithms such IIDMOT, IDMOT, etc.) to fuse later the information using fixed schemas and rules. This paper contributed a solid research on different algorithms and mathematical foundations to solve the different challenges presented. To end the paper, the authors presented some testing scenarios and experiments that validated the feasibility of their approach and to show interesting results.

Regarding the papers that included software architectures to support Human-Robot Interaction, the following ones were included in that category [138, 139, 144, 158, 160, 161]. In [138], Kim, Choi, and Kim presented an interactive software architecture for service robots that aimed to manage the robot's interaction with the environment and especially with humans. This software architecture was based on different modules governed by a "planner" module that manages the different robot motions. The software architecture has been tested in a tour-guide robot that operates at the National Science Museum of Korea. In [139], Parsons, Warner, White, and Gill presented an adaptable user interface and a software architecture to control a robotic arm. In this work, the authors presented how they designed a software architecture that enabled user interaction, interpreting it using software entities such as a dynamic data exchange and a dialog manager to later transfer the interaction to a robotic arm employed in users' rehabilitation. In this paper, the authors reported initial experiments with users solving different tasks with the system to validate the approach. In the paper [144], Mayer *et al.* presented a paper focused on users' interaction with industrial robots and industrial cognitive systems. They defined how users should interact with the industrial machines to improve the previously-achieved performance and how the interaction should be gathered and processed by a software architecture that plans the different steps, merges information and takes into account the different information from the users' environment to communicate the proper inputs to the industrial effectors. The authors reported how they applied this software architecture in different industrial facilities and used different user interfaces as well as how it affected the performance in the fabrication process and performance of different examples. In another work, Moshkina *et al.* [158] presented the TAME framework and software architecture, that tries to provide different user experiences to the users depending on the moment, users' attitude, mood, and emotions, in the context of humanoid robots and robots for leisure. In this case, the software proposal included different modules intended for the analysis of psychological aspects from the users, such as emotions, affective attitude, etc., to adapt the robot's behavior to them. The tests presented in the paper were related to how users should react to different users and what kind of behavior they should agree upon. In [160], Srinivasan, Murphy, and Bethel also delved into how software architectures could help in Human-Robot Interaction through the recognition of interaction such as gaze recognition. In this case, the authors presented a software architecture that recognizes physical stimuli from users and classifies them as input for robots to later respond to them coherently. In this case, the authors reported some results on how the analysis of the inputs from users are recognized and processed. Lastly, in [161] Veloso, Filho, and Barreto presented a middleware for robotic applications based on a software architecture based on web services that combine inputs from several possible interfaces to transmit the recognition to the different clients (the different robotic applications that respond to the user). In the paper, the authors presented the different schemas for recognizing interaction, and they presented different issues and challenges addressed by their proposal as well. As for the results, the authors included some interesting pilot experiences using interactions from users to make robots respond to them, etc.

Related to the other two previous categories, the papers that proposed different architectural solutions to support human-machine interaction can be found in [130, 131, 134, 141, 152]. In this set of papers, different common issues addressed were observed, such as how to support the interaction of impaired people with hardware/software systems [131, 134] or how to design adaptive interfaces in the context of human-machine interaction [130, 141]. The paper by Pires, Honorio, Lopes, Nunes, and Almeida [131] presented a quasi-intelligent wheelchair that relies on a software architecture to provide support to the impaired user. In this case, the paper described the different main modules that comprised the software architecture, the GUI designed to interact with the user and some different modules that introduced some features related to understanding and using other natural interfaces such as speech or how the wheelchair could adapt its control based on the user's behavior. Despite it being concise, the paper included a minimal practical experience using the wheelchair and showed how users could take advantage of the different features offered. Similar to this work, the paper by Caruso *et al.* [134], included a research work on how a software architecture could include different interfaces and characteristics to enable impaired users to use an assistive system. In the paper, the authors included interfaces such as BCIs or tablets as the main input aids for people with disabilities. The paper includes a full software architecture that included different assistive technologies and approaches to help these people to utilize different environments such as smart homes or to help them in their daily routines or tasks. In this paper, the authors exhibited different pilot tests on how users utilize their proposals and how these could help in their daily lives. As for the papers that dealt with the designing of adaptive interfaces in the context of human-machine interaction, the paper by Bozzon, Brambilla, Fraternali, Speroni, and Toffetti [130] described how an architecture proposal called MyHMI could help in adapting the HMI processes to the users' desires and preferences. The architecture is composed by different modules that deal with different interfaces and physical devices as well as other modules that analyze the signals and inputs gathered to unlock the proper responses from the machines that act as effectors. This custom treatment of the inputs and control over all the processes enables the architecture to adapt or customize some components and GUIs to meet users' preferences and characteristics. In another work, [141], Heigemeyr and Harrer presented similar work, regarding adaptability, but in the case of automotive human-machine interfaces. The work described in the paper was based on how the different interfaces available in vehicles to achieve better performance in tasks related to bidirectional communication and information between the vehicle and the driver can be adapted. To do this, the authors commented on different software components that analyze the inputs from the users and the vehicle's current status to enhance the communication and to adapt some elements such as dashboards or its behavior, etc. The last paper in this category is [152], by Tessadori, Bisio, Martinoia, and Chiappalone. This paper presented how users could interact with BCI interfaces, describing how the brain waves are captured and how a software architecture could enable the better understanding of the users' orders and actions and how they could be translated to other devices to perform the desired tasks. Besides the HMI-specific content, the paper also focused on the different brain waves detected and other issues and challenges presented by this kind of physical interfaces.

The next group of papers to be discussed are those related to software architectures to support systems' context-awareness. In this group, 4 papers were found [136, 145, 148, 159]. These papers could be separated into papers that present contents related to context-awareness in smart environments [145, 148], papers that included context-awareness to access the ubiquitous Internet [159] or papers that included context-awareness in mobile applications to enhance communication with user [136]. For context-awareness in smart environments, [145], de Alencar, Machado, de Oliveira Neris and de Almeida Neris presented a low-cost software architecture and system in their conference paper that is capable of modeling users' profile using ontologies and providing adaptable mobile user interfaces depending on users' conditions. The work by Vega-Barbas, Pau, Martín-Ruiz, and Seoane ([148]), in contrast, was focused on a software architecture that enables context-awareness in assistive and smart spaces. In their approach, the authors presented a way to articulate the interaction between users (patients at home or hospital, or others) and the smart space (their home, the hospital in the case of patients) by using different interfaces and devices and how they could use the different communication schemas defined to transmit their needs to their caregivers or assistive systems. Also, the authors presented other possible implementation scenarios. In this paper, the authors discussed all the levels of interaction and communication, providing a full overview of the proposal and its implementation in real-world scenarios. Related to providing context-awareness access to the ubiquitous Internet, Malandrino, Mazzoni, Riboni, Bettini, Colajanni, and Scarano presented in their paper [159] a distributed framework called MIMOSA that relies on a software architecture that manages the adaptation of web contents to the heterogeneity of networks and client devices as well as to the variation in users' environments. In this software, several modules received information from clients, from the environment, reasoned the context to plan the awareness and adapted modules and services provided to the users. The paper intensely discussed all the different issues solved and challenges addressed, and to be addressed regarding the user, the communication between components, the adaptation of contents, etc. As a final contribution in this paper, the authors reported on different tests and metrics related to MIMOSA, such as adapted pages response time, etc. To finish with this group of papers, which are related to context-awareness, the research by Pittarello [136] was found. The author describes a conceptual model and a software architecture on how to provide context-awareness to engage emotional users with mobile applications. To pursue this goal, the software architecture and conceptual model worked with different inputs from the user (the interaction, information from sensors, etc.) that allowed for the adaptation to different views in mobile applications and websites. The author presented some preliminary results in a case study related to a website that narrated different aspects and stories about San Servolo island, with 6 users who had a background in Fine Arts.

As for the papers that defined software agents to adapt software architectures to different contexts, can be highlighted the following ones [137, 150, 163]. From these 3 papers, there were 2 from the same authors (Jacquet, Bellik, and Bourda) [137, 163] that were part of the same research study (one is a continuation of the other). These two papers were related with a software architecture and model called KUP (Knowledge sources, Users and Presentation devices). In [137] Jacquet, Bellik, and Bourda presented the central concepts and foundations of KUP model. The architecture is based on a taxonomy of interaction modalities and user's profiles that are used to select the main interaction source available and how to use it to collaborate with the software architecture to solve different tasks in software or smart environments. Also, in [163] the authors used this basis to present more interesting case studies and scenarios, as well as to provide richer reports about the results using KUP in real environments and how it could be used, for example, in dynamic display of information for mobile users, or with natural interfaces deployed in smart environments. The last paper in this subset is [150], by Dimakis *et al.*, that described a framework and software architecture that can be used to facilitate human-centric services. In this case, the software architecture was divided into 3 different tiers: one that uses sensors to acquire context from users, another multi-agent tier that models and tracks higher level contextual situations and the last one that includes actuating services. The software architecture uses this agents-based schema to manage different multi-modal inputs, to provide remote participation features between different users, to enable surveillance systems, etc. That is, the authors provide many different scenarios in which to employ their approach and some metrics on case studies to attempt at a demonstration about its capabilities and features as a generalist solution.

Related to the papers focused on improving usability of systems [17, 149, 151, 153, 164], the following should be highlighted: In general, the papers found in this category were mainly theoretical proposals and studies related to the software engineering behind the improvement of usability using software architectures, especially in mobile devices and mobile contexts. In this group, however, two papers were not focused on mobile contexts. These papers were [17], by Bass, John and Kates and [153], by Folmer and Bosch. These publications were focused on common software patterns that could help to improve usability when using GUIs in computers or other devices, but bearing in mind that they could be applied to other contexts. To propose the different patterns, the authors investigated different interaction processes and their relation to well-known options and features available in modern software (actions and options such as undo, redo, etc.). With a different aim, the authors Biel, Grill and Gruhn [149], Seffah, Mohamed, Habieb-Mammar and Abran [151] and Capilla, Carvajal and Lin [164] specialized their research on the mobile context, proposing different software patterns that responded to usability challenges that were raised in these environments. In these papers, the authors researched different usages of mobile software, on the tasks performed by users and available aids to the users, to formalize how the most common ones could be enhanced, thus improving the user experience and usability in general. The principal value of these papers was that they contrasted with most of the proposals retrieved in this review, as they tried to formalize and propose standard solutions to common problems, avoiding *ad-hoc* solutions to problems that challenge users.

As for the remaining papers that were not grouped, their contents are commented below: in [129], Mackin, Gonia, and Lombay-Gonzalez presented a software architecture designed to connect hardware and software efficiently, to be used by astronauts during their training. This connection was not intended to exchange information between different hardware or software components or even with the user, but to enable new efficient interaction manners between users and devices in a special context such as Desert RATS field-testing. The devices presented in the case study simulated those that astronauts have used in space missions and presented some issues and challenges that were inherently different in the case of goals, interaction or tasks. In the case of [146], Rego, Moreira, and Reis defined interaction models for the use of 3D contents in serious games. In their paper, the authors proposed several natural and multimodal user interfaces in rehabilitation, similar to those offered by different devices such as Kinect, eye gaze trackers, joysticks, etc. Using these devices and the natural interaction that is linked to them, researchers defined a system that tracks and aids users in home rehabilitation for different cognitive and motor disabilities, implementing different strategies related to serious games to foster the users' motivation. As for the architecture, the authors proposed a layered architecture with different modules within each layer. Each component is responsible for different issues related to the gestures recognition, interaction, gamification, etc., and communicates with each other component or layer to collaborate in helping users to complete the rehabilitation tasks. The authors tested their approach and different usability tests with 20 healthy users, with some positive results (they did not include many details about this in the paper). In paper [132], Tiefenbacher, Bumberger, and Rigoll presented a system based on modular software architecture to support the evaluation of touch interaction in industrial environments. In the paper, the authors presented different touchable interfaces

(7'' and 12'' screens) and demonstrated how they evaluated users' interaction with this kind of devices using a modular architecture that recorded the interaction with each software component and the different gestures related to the industrial activities and tasks. Moreover, the authors presented some interesting results using their system, which included users' performance with each screen size and different software interfaces and gestures, and they measured task completion times, clicks, and other scores using a SUS questionnaire as well. The paper [142], by Jalaliniya, Mardanbegi, Sintos, and Garcia described a system to support users' eye interaction with other systems. The system is based on a gaze tracker built using Android, and relies on a modular architecture that controls the different issues related to gaze tracking, computation of user inputs through the detection of eye movement and other events and communication with external systems in an efficient manner, enabling the proper features for simple and transparent interactions. The authors tested their approach on a Google Glass device and involved different users in testing different issues such as precision of gaze tracking, easiness of interaction, etc. In another work, Sutcliffe, Thew, and Jarvis presented a paper [162] that aimed to support the integration of software requirements and HCI. In this paper, which was more theoretical than the others in this set, the authors described how they experienced the integration of user-centered requirements engineering in the definition and development of complex systems. In this case, the integration of these user-centered requirements enabled the consecution of better interactive systems which are closer to the users' needs. To present the experience, they proposed different workflows intended to be performed by researchers and system analysts to integrate the user-centered requirements from the beginning of the system conception. In general, the authors reported a positive evaluation of a system and prototype developed using their approach and described how this kind of requirements techniques could be applied jointly to other design methodologies and methods such as storyboards, scenarios, and prototypes to achieve better products and receive positive feedback from the users. To finish the discussion of this group of papers, the last three papers were related to multimodal interaction: the conference paper [133], by Calandra, Caso, Cutugno, Origlia, and Rossi presented CoWME, a framework and software architecture designed to evaluate cognitive workload during multimodal interaction using different systems. This software architecture, in words of the authors, "takes as input an arbitrary number of input streams containing data reports about the cognitive workload estimated by dedicated modules, each concentrating on specific modalities. Typically, there will be one monitoring module for each available interaction modality". In the modular software architecture, each module is responsible for collecting and interpreting the observed cues concerning cognitive workload. Using these inputs gathered, the system analyses and made decisions to measure the cognitive workload experienced by the users effectively. The authors showed a deployment example of their system, which was used to analyze the users' cognitive workload in an Android device through recording touch interaction, speech, and mydriasis (pupil dilation). In this example, the researchers received information from 5 people, who played a game in an Android tablet. The results from the example showed that the platform accomplished the tasks for which it was designed. In [154], Marsic and Dorohonceanu described a software architecture that supported multimodal interaction for user collaboration. This software architecture articulated the combination of different interfaces (in a multimodal way) to enable users to collaborate more effectively with peers. In the case of the presented approach, the authors included algorithms for multimodal information fusion, integration with shared workspaces, dynamic customization of interfaces, etc. To test the proposal, the authors conducted some experiments with four teams (each one comprised by 3 participants) to perform several tasks using/without using different software components enabled by the software architecture. In two out of three tasks, they demonstrated that by using their software components, the users collaborating in solving the tasks obtained better results and performance. Moreover, lastly, in the last paper [156], Olmedo, Escudero, and Cardeñoso presented a software architecture that supported multimodal interaction with 3D environments (mainly virtual worlds). By observing the previous and different issues and challenges, the authors defined a specification language for virtual spaces with multimodal interaction or a software architecture to interpret the different multimodal interactions. As the main results of their paper, the researchers defined the XMMVR language to define the multimodal interaction spaces, created a modular software architecture that uses the XMMVR language to gather and analyze the multimodal interaction specified and evaluated the results obtained in a case study. In these results, the authors measured some variables from the users related to the multimodal interaction, such as the ease of use, satisfaction or ease of learning. In this case study, the multimodal interaction was very limited to typical interaction devices (mouse, keyboard, etc.) and was more intended to evaluate the language and software architecture developed.

Finally, should be noted some brief comments about the papers analyzed, in line with previous ones that appeared in this paper. In general, the software architectures proposed were not based on traditional approaches or well-known proposals. Most papers used custom solutions that were not designed to be reusable or generic for some purposes. The exceptions were: the paper [140] that based their proposal on the PAC-Amodeus [166], the chapter by [163] that was based on their previous work on the KUP model for multimodal interaction [137] and the work by [147] that used the MARIA model-based framework [167].

The dominance of research papers related to multimodal interaction reflected that gathering, analyzing and interpreting several inputs or interaction sources in real time or as a whole is still a problem for many systems and interactive environments. It was detected that had been published a massive amount of work in this area, but there is still room for more improvements in the combination of input sources through the enhancement of information fusion algorithms, multi-input, and multi-purpose architectures, etc. For the HCI/HMI itself, was observed a lack of interaction analysis as the primary purpose. In the case of theoretical papers, they approached this problem through the use of patterns, but they did not provide real-time solutions to enhance usability or define adaptability to the users' profiles or behaviors. Related to the techniques found, the usage of AI-related techniques or advanced statistical methods in this subject is missed. Only 3 papers use this kind of approaches to analyzing the interaction; the other few works found dealt with the analysis based mainly on heuristic rules. As shown in other application fields, that modern data-driven or AI-driven approaches could enhance the research and provide better results. Also, in other cases, the interaction analysis is not present as it should. In this sense, exists an important gap between academia and the pure engineering/industry work in this area. Many approaches followed in the industry nowadays are related to analyze the users' interactions within their different systems to provide proper feedback, to analyze behavior to ad-targeting users, etc. It is possible to agree that some of the goals pursued by the industry are not aligned with those from the academia, but the academia could provide new valuable approaches to make this kind of analysis better, more efficient, more responsible to the user and ethically adequate, etc.

Regarding the communication between components/layers and other systems, can be highlighted that most systems used typical data structures encapsulated in files such as XML, CSV or JSON, and not many proposals relied on web services, APIs or microservices as in other research areas. Lastly, should be remarked that many papers did not provide any kind of —or proper— evaluation or assessment of the support or analysis provided by the software architecture; they only included brief reports about tests or minimal experiments/experiences with users.

2.6 Threats to validity of this systematic review

This systematic literature review and mapping aimed to identify, categorize and analyze the support provided by software architectures to HCI/HMI processes. As with any research method, there could be threats to its validity and limitations. As other authors have mentioned [121], in this kind of mapping and literature reviews, the results presented "may have been influenced by certain uncontrollable limitations". One of the main threats to the validity of this kind of investigation is the author's bias in certain aspects in papers or subjects. In this case, some strategies (such as the quality assurance checklist [115]) used to mitigate this bias were applied. Also, the number of researchers involved in the selection of papers could be considered a bias. To reduce this bias, the Ph.D. candidate was the primary reviewer, and the two supervisors reproduced each step taken in the whole process to ensure its validity. Also, the different resources provided (such as the spreadsheets or the public GIT repository [120] that contained the results from databases as well as other resources used in this research) were intended to reduce or remove any bias and make the whole process reproducible for the reviewers and readers. In the case of the search string, the author previously evaluated some papers to monitor the results and validate the search string (i.e., that papers appeared in the results gathered from the different databases used), so it could be assumed as a kind of evidence to ensure the query correctness. Regarding the databases used, all the most relevant digital libraries and databases in the field of computer science were included. Thus, one of the main typical threats to validity that usually arise in systematic literature reviews was eliminated. Other databases such as Google Scholar could have been included in this research study, only considered those that ensured contrasted quality in their contents were considered.

3 Experimental research

This chapter presents the different incremental experiences and experiments performed during the research to try and answer the initial and central question posed in the introduction section of this manuscript.

The chosen cases studies, based on the scenarios proposed in section 1.2, are the following:

- Scenario 1: Highly interactive scenarios where the users should develop individual activities or tasks based only on their interaction without human aid. As presented in section 3.1, the case study selected for this scenario is an educational environment in a virtual world. In this case, the environment is the Usalpharma island hosted in Second Life. This island is designed and built to serve as an educational facility to teach and train students in Pharmacy.
- Scenario 2: Systems including a large number of users collaborating or using concurrently the same resources to solve similar tasks. In this scenario, other stakeholders evaluate the users' performance in solving their tasks. The case study selected is a MOOC that includes users' interaction with learning contents and conversations among them in a web platform and social networks. The experiments performed and results achieved using this case study are discussed in section 3.2.
- Scenario 3: Information-intensive applications where the users could need aid and be engaged in solving appropriately the challenges proposed (with challenge understood as the need for solving a task). The selected case study is an extensive web form (consisting of between 30-90 questions) owned by the Spanish Observatory for University Employability and Employment. The study pursues to gain knowledge on how to improve the users' experience and engage them and improving their performance in completing the form. The study, as well as the technical solution provided, are presented in section 3.3.
- Scenario 4: Complex systems where the user might not have enough previous knowledge to solve a task efficiently. The case study selected

for this scenario is devoted to understanding how to aid programmers when they develop code in a challenging and new context as quantum programming. In this case, the research was developed by the author during his predoctoral research stay (and is still in progress) in collaboration with the IBM Research AI & Q team. The experience and the preliminary results relevant for this thesis are shown in section 3.4.

Figure 7 explains the sequence followed in the research using the different case studies and scenarios.

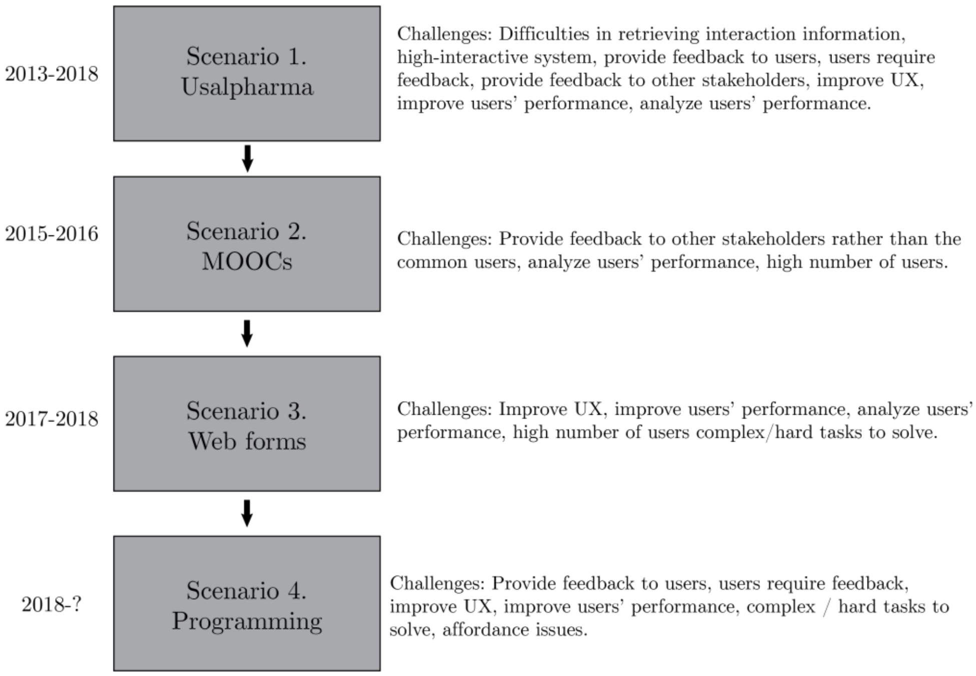


Figure 7. Planification of the experiments carried out during the thesis and challenges present in each scenario.

In Table 14 the overall differences and common points between the case studies regarding the HCI-related challenges present in each of them are explained.

The next four subsections present each case study representing one of the four scenarios selected for this thesis. This pave the way to try and answer the questions posed in section 1.2. Each subsection outlines the context, goals, materials and methods, results and discussion per each case study. Later, in section 4, a global discussion is offered with the intent of answering the overall research questions.

Table 14. Differences and similarities between the different study cases.

Challenge	Scenario 1: Usalpharma	Scenario 2: MOOCs	Scenario 3: Large forms	Scenario 4: Programming QC
Difficulties in retrieving interaction information	X			
High-interactive system	X			
Provide feedback to users	X			X
Users require feedback	X			X
Provide feedback to other stakeholders beyond than the common users	X	X		
Improve UX	X		X	X
Improve users' performance	X		X	X
Analyze users' performance	X	X	X	
A large number of users		X	X	
Complex/hard tasks to solve			X	X
Affordance issues				X

3.1 Improving virtual worlds users' experience and giving feedback on their interaction to different stakeholders

There are many definitions of what constitutes a virtual world and what its primary characteristics are. According to Schroeder [168, 169], a virtual world environment is "[a] computer-generated display that allows or compels the user (or users) to have a sense of being present in an environment other than the one they are actually in, and to interact with that environment". In a Virtual World, the user creates a virtual alter ego or avatar, representing him in a simulated space. Through avatars, the users can interact with the 3D environment, other users or other kinds of contents [170]. For example, some uses of Virtual Worlds with regard to eLearning [171], playing [172], marketing or e-commerce [173], socializing [174], etc.

The most relevant results of the research carried out using this virtual world have been published in [175-179]. Also, the papers are available in Appendixes 6.1, 6.2, 6.3, and 6.4).

3.1.1 Context

User perception and engagement with technology is a crucial aspect of a successful technological system, regardless of application. Information about user's perception and engagement can be used by system developers, administrators, designers or managers to adapt, enhance or modify the technology to improve user acceptability, interest, feedback or performance. This type of research and analysis has a significant presence in the field of eLearning. Within this field of online learning, behavioral and engagement analysis could be used to propose, implement and correct various types of platforms that enable education and skills acquisition via the use of Internet-based technologies. Many authors [180-182] have explored the field of behavior analysis and engagement related to online learning environments as a tool to determine how users feel and use technology to obtain insights into the acceptance and adaptation of these eLearning ecosystems in the Learning Process and increase success.

Tracing behavior patterns and measuring engagement on these platforms enables the determination of user interest in specific features or content and whether these features are exploited properly. These measurements enable platform managers to make decisions, correct unexpected uses, promote specific content, perform actions to avoid dropouts, improve system adoption, and adapt the structures or content of eLearning platforms to users.

An educational environment appropriate for this type of study is an Educational Virtual World. This type of environment provides an exciting field of study of user behavior and engagement because it is based on user interaction with a 3D environment and with other users. This feature, combined with the representation of the user via a virtual alter ego and features such as text-based chats, voice-based chats or movement between different islands or lands, allows users to express themselves very different compared to other educational environments [183, 184]. Thus, user preferences and desires in the field of Learning are often reflected more accurately than in other eLearning ecosystems.

Since the beginning of the use of Virtual Worlds in educational contexts, several authors have pointed out the problems related to knowing what is happening within the 3D environment about users' behavior, the usage of the virtual world features, etc. [185-188]. Among the goals of this knowledge of what happens in virtual worlds, highlights the possibility of evaluating offline (outside the Virtual World) what happens online (inside the Virtual World). Achieving this goal, teachers related to Educational Virtual Worlds could know completely the learning process that takes place within it so that they could measure the learners' evolution, interest, contents learned, results they get, or even predict abnormal behavior, drop out of students, etc.

From students point of view, the data retrieved from their usage of a virtual world could help to improve their learning process. For example, if there is a system integrated with a virtual world that guides students to achieve learning goals [189, 190], discovering new training process, warning them if they fail, scoring their progress and achievements in educational activities, etc. like a virtual master who helps them improve. The presence of a supervisory system which analyzes what happens in the virtual world, and can give results and answers to in-world events, can facilitate the implementation of training scenarios with a fixed number of paths or solutions so that the system itself could assess achievement of the different phases or solutions. Thereby, the digital supervisor could help automate the educational process or professional training that occurs within the virtual world, facilitating independent learning and repetition of closed cycles of knowledge acquisition by students.

During the thesis and the previous years, the author was involved in the Usalpharma project and other educational initiatives involving Virtual Worlds. During his participation in these projects, the author developed scripts to extend the functionalities and features offered in the virtual laboratory and allowing students to receive better feedback on their actions and improve the immersive experience within the virtual environment. Giving the fit between the objectives of the first scenario and the author's experience and the features requirements of these virtual worlds, they were selected to be a centric part of the research.

3.1.2 Goals

As outlined in section 1.2, the objectives pursued in this scenario are:

1. Collect and analyze individual users' interaction from a highly interactive environment to extract knowledge and evaluate users' performance solving tasks.
2. Test how to take advantage of the knowledge gained, from a software perspective, to feedback users and improve their results.

Considering the case study Usalpharma, its context and the inherent challenges of a virtual world like Second Life, the specific goals can be extended to:

1. Since Second Life does not offer reports on users' interaction (as will be explained below), create a system that collects this interaction.
2. Analyze the users' interaction collected using this tool and determine what kind of users' behaviors and performance indicators can be discovered using the data gathered.
3. Offer feedback related to the interaction to the students that use Usalpharma facilities to learn about Pharmacy.

4. Offer feedback related to the interaction to the teachers that use Usalpharma facilities to teach about Pharmacy.

3.1.3 Materials and methods

Materials and methods related to developing software that collects information from Second Life about users' interaction and to give feedback to stakeholders involved in teaching/learning using virtual worlds

In the case of this scenario, the work has been performed using the Usalpharma island within the Virtual World named Second Life^{TM} (SL). Second Life (SL), shares the same features than other virtual worlds: is a virtual space available on the Internet in which it is possible to simulate real-life situations and scenarios, it is an entirely immersive 3D environment in which everything is created by the users or residents who interact with each other. Also, SL has a powerful tool for the creation and edition of objects that allows the construction of buildings, devices, furniture, and clothes. In these, it is possible to insert programs called *scripts* which endow them with functionality, mobility, lighting, etc. All this makes SL an exciting learning tool [191-193] that is widely used in the health sciences [194-201] although the regular appearance of other virtual worlds has removed his supremacy in other fields. Moreover, SL (as well as other private virtual worlds) does not provide APIs or specific methods for interaction and exchange of information with other software platforms; they only provide in-world APIs to connect with web services or retrieve information from external websites.

On the Usalpharma island hosted in SL, a property of the Department of Pharmacy and Pharmaceutical Technology from the University of Salamanca, teachers, and students have been carrying out activities to develop professional skills in undergraduate and postgraduate students of pharmacy since 2010 [175, 176, 178, 179, 202-206]. These activities take place in various virtual facilities such as a community pharmacy or a laboratory for drugs quality control.

The most relevant virtual facility for this research is the Usalpharma Lab. This laboratory simulates the installations, equipment, documentation, and tools like a real quality control laboratory in the pharmaceutical industry that fulfills the legal regulations and requirement [178, 205]. This kind of installations in the real world are very costly and unusual in universities. Usalpharma Lab has been used for training of pharmacy postgraduates in quality assurance. To carry out these practices, both avatars of teachers and students met initially in the virtual laboratory. The teacher guided and evaluated the students during the activity. However, in words of the teachers and technical staff involved in the project: "it would be ideal that the students could access the laboratory whenever they wanted, using it as a tool for self-training". So, data generated through the student activity could be used by the teacher to control and evaluate their activities, without having to be present at the same time. These data could provide information of high relevance for the design of new practical activities as well as the evaluation and monitoring of the correct implementation of them [91].

Materials and methods to analyze data and detect users' behavior and performance metrics

To develop this part of the research work was used an Educational Virtual World developed by the University of Salamanca. This Virtual World, called USALSIM [89, 207], is designed to provide virtual practices and immersive learning experiences through a 3D environment. In 2012, a series of 3D environments permitting the development of professional practices and learning in the following areas of knowledge were incorporated into USALSIM: Pharmacy, Biology and Biotechnology, Law, Humanities, and Chemistry. The type of activity varies depending on the specific needs of each area of knowledge and practice; some require personal work by each user (purely interacting with 3D scenarios inside the Virtual World), while others require more discussion or role-playing tasks (interaction with other users).

This virtual world was developed under the platform known as OpenSimulator or OpenSim [208]. This open-source platform allows the managers of the Virtual World (in this case, developers at the University of Salamanca) to control all data related to users and the 3D environment, enabling a thorough analysis of the selected data. In contrast, other Virtual Worlds (like Second Life for example) do not allow the retrieval and analysis of detail in such detail.

Microsoft Excel was used to extract raw data (only organized in parameters and not modified) from the Virtual World for exploratory analysis. Using this software, the researcher performed a manual analysis based on clustering of standard features and measures of users and the use of basic statistics to analyze the measures. Automatic methods were not used for the analysis because this study was primarily an exploratory analysis and was intended not to address pre-defined questions but to develop findings related to the research goals.

For this study, data representing different characteristics and areas of usage were retrieved from the USALSIM Educational Virtual World and analyzed separately and in combination to identify user interests, behavior patterns or engagement measures.

As many authors have noted, Virtual Worlds feature a robust social use component [209, 210], and even in those Virtual Worlds that implement learning activities, users often interact with each other, whether required by the educational task (such as role-playing activities) or not.

The data used were collected in two months, between November and December 2012, during testing before deploying this Virtual World to the general student population at the University of Salamanca. These preliminary tests were intended to study student acceptance of the Virtual World [211], its relevance to education in the University and its effectiveness for skills acquisition and content learning. After the tests, both teachers and students assessed aspects of appropriateness and acceptability through surveys [207]. For this part of the research, the author does not intend to evaluate the acceptance or effectiveness of the platform for learning or its specific use by users during the test months, which could be determined using surveys after the pilots or based on student grades. The author does intend to explore user behavior and usage patterns to obtain more in-depth knowledge and explore the value of this knowledge in decision-making processes about the learning process in this Virtual World.

The tests involved 75 users, and data were collected on various aspects that might indicate usage characteristics, behavior and engagement indicators in the Virtual World. These 75 users were classified as follows:

- One system administrator: This user was not included in the collection and analysis of data because the number of actions performed in the 3D environment and the time used in the system introduced too much noise in the data set.

- Ten teachers: These teachers were labeled differently in the analysis to identify potential differences in behavior between teachers and students.
- Sixty-five students: Students in the various knowledge areas described above (Pharmacy, Biology and Biotechnology, Law, Humanities, and Chemistry) volunteered to try this new learning environment at the University. Because the testing stage was more focused on knowledge about use and engagement, the analysis did not differentiate among the knowledge areas of the students.

Data concerning four of the key features and options in Educational Virtual Worlds [91] were retrieved: voice chat-based features, text communication-based features (between users, messages between objects, etc.), session information (time in each, total number, average time, etc.), and movement inside the Virtual Worlds (between different lands, islands, etc.). These features were chosen for analysis because they provide a range of information about users' desires, interests, usage patterns, or habits in both Educational Virtual Worlds and other eLearning Systems [180, 211, 212]:

- Voice API calls: Data on the number of times the 3D characters used the voice service within the Virtual World. These data provide information about communication only between users and can reveal specific social characteristics [209, 213].
- Messages with objects and users: Data on the number of times that 3D characters made use of private messages to each other or received information from the different objects in the virtual environment. These data provide information on both social uses, such as the 3D scanning environment and the use of educational resources [214].
- Sessions: Data about the time spent by users within the Virtual World and its distribution in time, number of sessions, average session duration, etc.
- Movements: Data about the movement patterns of users. The number of movements may be an indicator of engagement, while a movement to a specific territory may indicate interest in that territory [184, 215].

Data about other indicators, such as clicks in 3D contents or objects, text discussion themes, changes in avatar appearance, and profile modifications and descriptions, were discarded because they are more difficult to measure in an initial exploratory analysis of this type of platform.

To identify patterns, understand how users use the environment, and determine their involvement and engagement with the platform, an exploratory analysis of the data was performed. A detailed exploration of the data was performed to reveal trends and relevant issues in the collected set of interaction-evidence. A panel of experts in the creation and use of learning scenarios in Educational Virtual Worlds was interviewed to obtain some general heuristic rules and their application and validation or rejection based on trends and relevant data.

The expert panel comprised five professors from the University of Salamanca, none of whom was the author of this thesis or interfered actively in the realization of this research study beyond the discussion of patterns or behaviors. The experts had five years of experience in the definition and use of Educational Virtual Worlds in a Higher Education context, specifically in Second Life, which was launched in 2003, and thus their experience encompasses nearly half the time this system has been available on the market.

Individual interviews were conducted with the members of the expert panel to achieve two goals: first, to establish some heuristic rules and widespread occurrences observed in their experience; second, to review each of the most characteristic patterns and rules obtained in the exploratory analysis. Conclusions and perspectives were also collected as part of the accomplishment of these goals.

This part of the section describes and explain the general heuristic rules disclosed by the experts; these heuristic rules helped the author to explain the patterns and association rules between the measures. Among these general heuristic rules, the following can be highlighted:

1. The experts first insisted that different types of users can be highly differentiated regarding the ease of use and engagement toward these platforms. Users (students or teachers) can be classified into three groups: those who reject this technology, those who find it difficult to

use but are able to employ it directly but effectively with practice, and those that operate seamlessly in these environments, display keen interest in it and are intensively engaged with the virtual environment. The experts also emphasized that this classification is often not reflected in the scores of activities conducted within the Educational Virtual World, although groups that use these tools more proficiently and with greater engagement often used more resources and achieved better results.

2. Consistent with the previous heuristic rule, the experts emphasized that users who spent more time on the learning scenarios (not in the Educational Virtual World, but in a particular learning scenario) tend to perform better in the learning activities.
3. Moreover, the experts stated that users without extensive experience in the use of tools and 3D (Virtual Worlds, video games, etc.) often refuse or are reluctant to use the virtual environment, occasionally even resulting in dropouts from the activity scenarios. These users may also underperform because of their non-acceptance of technology.
4. The types of features and resources used to depend on the learning activity performed. Experts agreed that users use more 3D objects, text-based chat or voice chat depending on the learning practice; for example, in practices with virtual face-to-face meetings, user tend to use chat (both text- or voice-based) more, while in autonomous learning activities, users tend to interact more with virtual objects.

As explained previously, engagement is a crucial factor in the successful adoption of a learning ecosystem by students, but one question must first be addressed: *What is engagement?*

Many authors have attempted to define engagement in the context of online learning environments [180, 182, 216, 217] using different factors, either concerning effectiveness, the element usage rate, platform acceptance and use, motivation, etc.

To clarify the engagement concept in a simple definition, it can be asserted that *engagement is related to the involvement of students in their learning process or tools*. This involvement is evident in many ways, such as the time users spend learning, the number of activities undertaken, the relationship between results quantity and quality, etc. [218].

In general, engagement factors are those characteristics that can induce a particular or superior interest in its learning cycle [95]. In the case of Virtual Worlds, for example, engagement features are those factors related to the interactions of users with both the 3D environment and other users that increase use and acceptance of the environment, thereby achieving better final results in the learning process.

As explained previously, the measured factors for this research work are the following: data about voice API calls, data about messages with objects and users, data about sessions, and data about movements. As stated previously by others, the interaction between the user and the environment is more reliable in Virtual Worlds than in other eLearning ecosystems because users identify their digital identity (the 3D character or avatar) as a replica or reliable representation of themselves [183, 184].

These factors all involve interaction with the environment and other users or the amount of time and the rate of use of the system or specific features of it.

Measuring the engagement factors reveals common factors, characteristics, behavior or usage patterns among users. As many authors have noted [183, 186, 212, 216, 219-221], these usage patterns and evidence of engagement can help explain how users engage with specific learning tools, enabling improvements or corrections of aspects that are not being used correctly, preventing dropout, and helping users to reach their learning goals [222].

3.1.4 Results

Results related to the development of software that collects information from Second Life about users' interaction and to give feedback to stakeholders involved in teaching/learning using virtual worlds

Between 2011 and 2013, the teaching activities carried out in the Usalpharma island and lab were under the direct supervision of teachers of the course and staff involved in the development of the 3D environments. That is, practice work used to be developed in the laboratory (in a single session and call for all students), was guided by a teacher and was supervised by the rest of the team, so that if a student had any questions or appreciation of the 3D scenario, content resources, possible variations or disagreements regarding present rules or any other aspect they could immediately make a query, getting instant feedback and sometimes very specific help. This stage, as far as results are concerned, was very positive for the initiative, students got good grades in their practice and showed satisfaction with the methodology and resources used in this virtual learning [206].

After this first stage in the implementation of this type of system to support teaching, the teaching innovation team wanted to go a step further, through implementing a system that would *replace* the teacher's guide during the students' practical sessions, controlling whether students were doing or not the practice within the virtual world, as well as to include a number of requirements on the feedback that students receive when participating in their practical work. That is, a system that allows taking advantage of the potential of virtual worlds in terms of the ability of autonomous learning, the possibility of not depending on the time and physical space (time, physical location of students) to leverage resources provided by virtual learning environments as well as presenting better immersive experiences to students within the 3D environment [193, 221, 223-225].

Before designing and developing a viable and optimal solution for this problem, researchers had to take into account a number of specific preliminary considerations of the problem to be solved (some of them have been outlined briefly previously), among which are highlighted:

- The problems posed by the interconnection of a private and closed system such as Second Life with any external platform. Not only does it present problems regarding integration, but there are also no public resources, logs, APIs or other tools to integrate with other systems [91, 226] beyond basic information on a number of users per day, etc. [227]. It presents additional problems because of the rigid characteristics regarding the ownership of the 3D objects, which hinders the usage for the described purpose. Even more, Second Life is provided "as is", which prevents any claims about failures, changes, etc. and requires users to modify the system behavior [228].

- The system to be developed as a solution must be able to coordinate the response to the whole problem, from data collection to the presentation of the same on any device or within the Virtual World. This implied drawing on well-known open technologies, standards that have an almost complete presence in every possible system. These standards must be present in the technological solution as well as communication protocols, data collection, etc. [229, 230].
- The technology that supports the management of processes, such as layer data management, must be able to withstand peaks of activity (many students interacting simultaneously in the 3D laboratory), possible massive load at specific time intervals (while carrying out the practical work) and concurrency in the interaction with the resources that provides performance without excessive erosion.

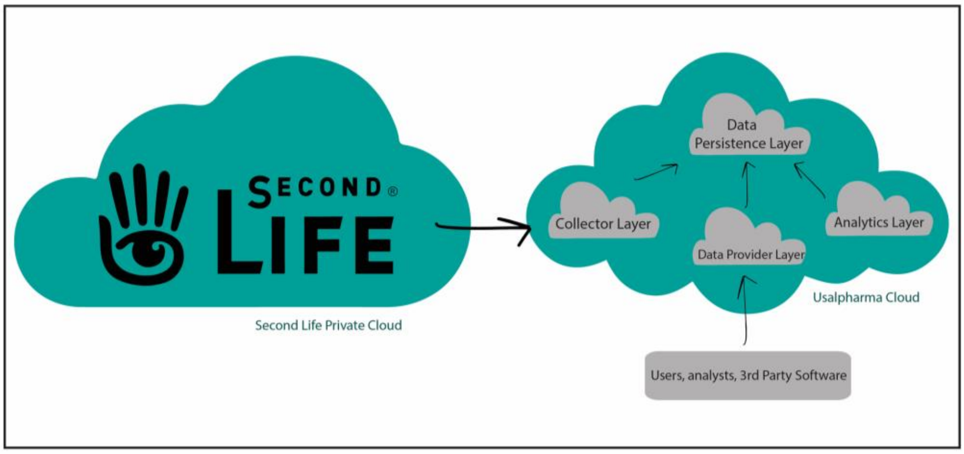


Figure 8. A simplified representation of the architecture of the proposed solution [226].

Since 2013, the author has been working to address the problem adequately, and in early 2014 a viable solution was finally proposed [226] (at least from a theoretical point of view) that provided an adequate framework and a proof of concept for the final development of a product. This solution could remedy the problem regarding integration of a system that monitors, analyzes, reports and helps the evaluation of teaching activities within the virtual laboratory Usalpharma Lab.

This solution is based on an architecture that could be deployed in a cloud environment. The architecture, based on the typical client-server schemes, consisted of a number of layers (data collection, data persistence of analysis, presentation, etc.) connected to each other in the same way (client-server depending on which requires the services of other interactions), enabling the deployment of each layer even in different clouds (with the potential to scale only those resources required, apply different technologies to each layer, etc.) due to the fact that the layers merely interact among them using services (Figure 8).

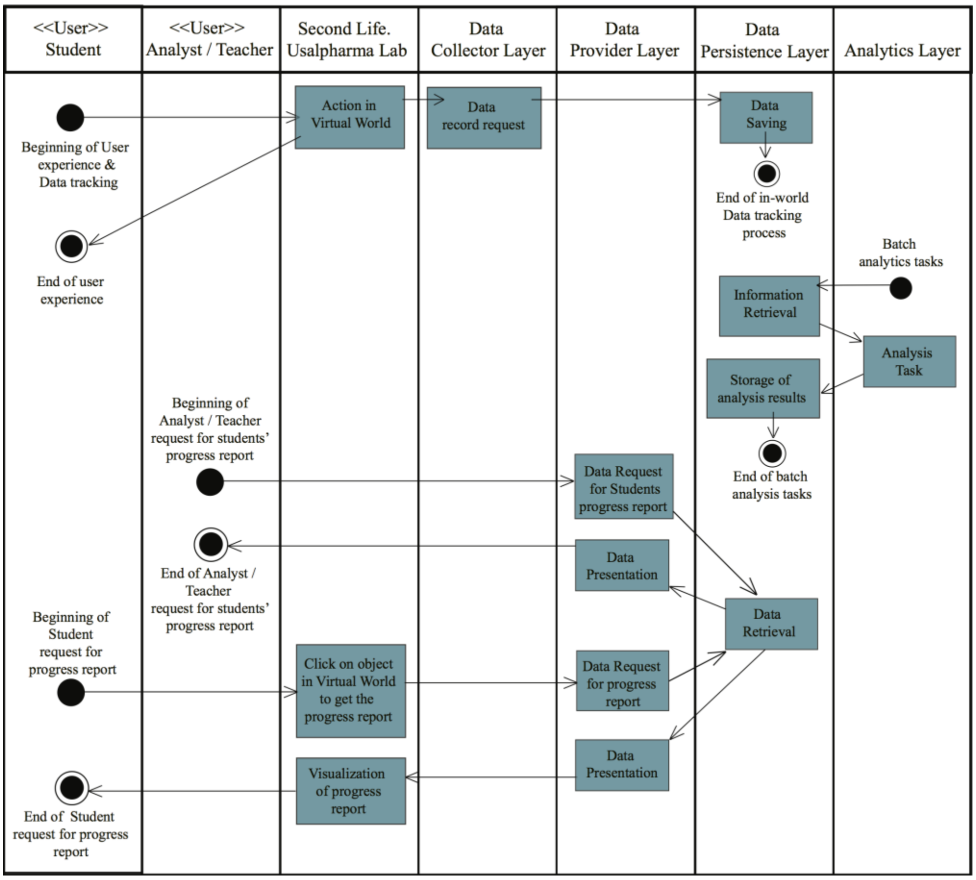


Figure 9. Activity diagram among users, Virtual World, and system. An updated version of the proposed in [226].

Regarding the standards and technologies used in this technical solution, the following can be outlined:

- *Software Architecture:* It is implemented using the web framework Django [231]. It consists of a series of applications that are responsible for collecting the Hypertext Transfer Protocol (HTTP) [232] requests that launches the Second Life client for recording data interaction, cleaning, and processing such information, contacting the persistence layer data and database MongoDB [233] data. Also, it responds to requests such as "display information" depending on the context/client from which they are requested (the same information is not served when the user launches requests from the web application or the Virtual World), calculating data associated with user interaction. Examples of these calculations are measures of the time spent by each student to complete their practical work [234], calculating the most relevant critical points achieved [176], reporting for each user or group of users, etc. These applications and the multiple layers composed architecture are responsible for maintaining the logic of the whole system as well as the bulk of the functionality.
- *Web Client:* From the Web client, teachers can consult all data concerning students and their work progress. Among the metrics to report should be available the number of collected interactions, the time spent by students while interacting with the 3D environment, the achievement by the students of the inspection of the various elements (review of safety measures laboratory, review of equipment to be audited, documentation, etc.) as well as comprehensive reports for each class of students, so that a teacher can know in a detailed way the performance of each student, or acquire an overview of the group of students in their practical activities. This web client is developed to be used from any device (PC, tablets, smartphones, etc.). Figure 11 presents some data visualization of various metrics that teachers can know about student performance in Usalpharma Lab.
- *Second Life client:* This client is somewhat different from the web client because it can be used by any avatar (user in Virtual World) that is in

the laboratory. The client is used through interaction (click) on an object in the laboratory, so that the student sees a dialog showing in real time the percentage of critical points assessed, both in the day when a query is made and, in general, making suggestions on what the student should re-audit. During the first year deployed, the system showed only feedback about the particular day of the request, without giving an overview of the development of the practical activities on different days. Figure 10 shows this type of activity report.

At the same time, the Second Life client has an even more essential role, and that is to send data to the system so that it registers every evidence of interaction between the user and the virtual laboratory (data entry to the platform). To do this, all 3D objects to be audited by students in the virtual laboratory have an associated script that throws an HTTP request to the application, which is responsible for collecting information about user interaction to record the action that is occurring at the moment. This request is carried out silently and transparently for the user; the object itself is responsible for collecting the data on the date and the exact moment of interaction (timestamp), with what user has interacted, where was located in the virtual lab, etc. The user is not aware of the transmission of the information to the system, although all students are previously notified of the monitoring carried out in the laboratory.

Regarding the standards and technologies proposed in this solution, the following can be outlined:

- Use of a (simplified) variation of the semantic standard Resource Description Framework (RDF) and its triples [32, 235] to transmit structured and rich information about the interaction in the Virtual World to the data collection layer.
- Use of the HTTP protocol [236] for the transmission of messages between the Virtual World and the layers of data collection and provision (in both directions, the Virtual World to the system, and the system to the Virtual World), explicitly using the part of the protocol concerning operations, requests, status codes, and headers.

- Intensive use of web technologies to return information to users (teachers and students) regardless of device, operating system or software (such as Second Life viewer).
- Using NoSQL document-oriented databases (MongoDB [233] specifically) that allow modification of data architecture transparently and without penalty arising from the persistence of the layer architecture (only possible penalties in the logical layer derived from changes in its own data models [177]), and maintaining good performance in information processing, and high scalability in production environments. [237].
- Use of server technologies that have demonstrated their capacity for data processing HTTP requests and adequate performance under massive demand. In this case, the technology proposed was the web framework Django [231], which develops robust solutions using Python and has many libraries of its own (and others of the language) to facilitate the development of stable and agile projects.

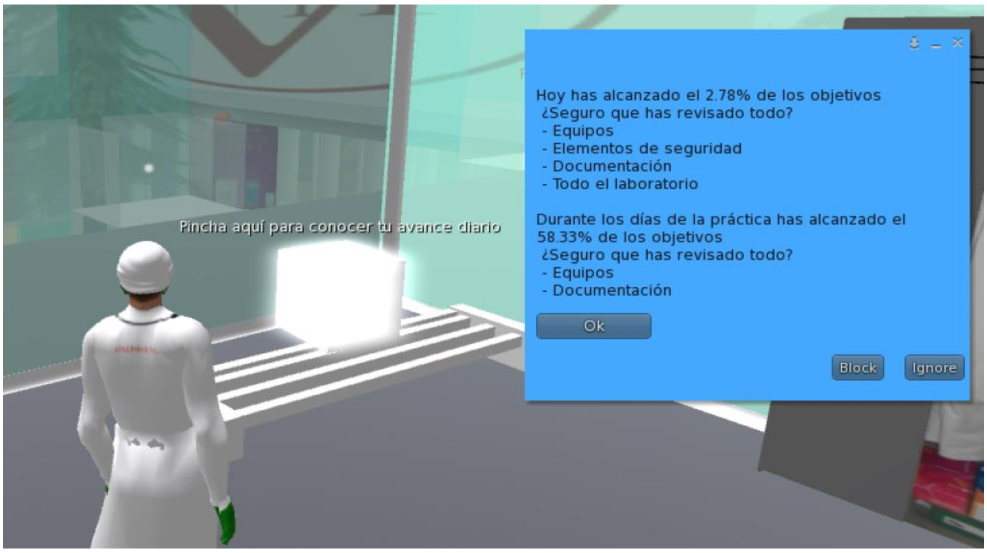


Figure 10. Real-time information that students can get in Second Life about the audit performed. In the Spanish text that appears in the figure, the system informs the student about the performance (% of the goals in the day and all the days of the practice) and suggests to the student the category of elements which lack interaction.

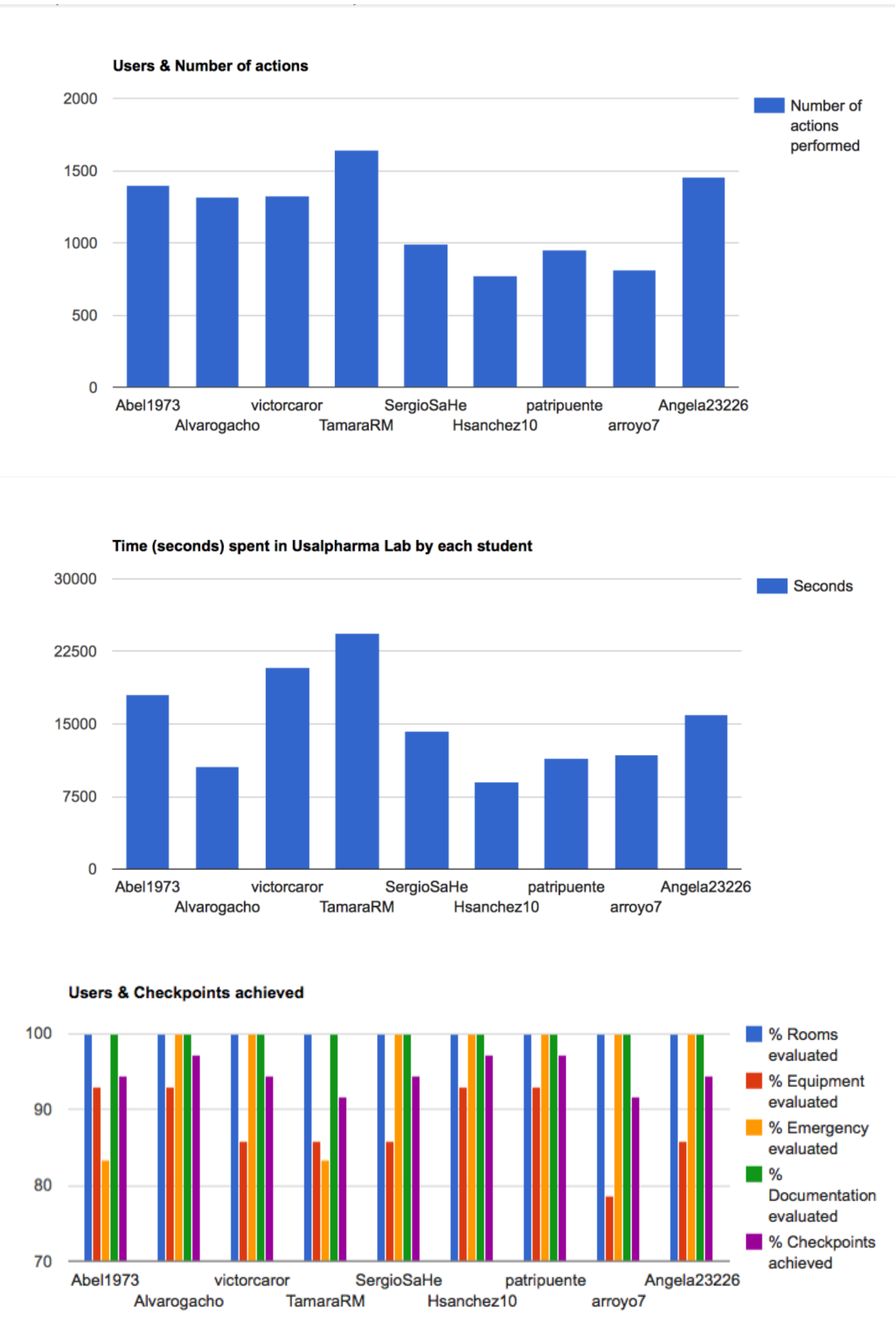


Figure 11. Some data visualization presented in the Web client. They show different metrics related to student's interactions in the practical activities during 2015.

This solution provides a flow of activity as proposed in the activity diagram shown in Figure 9, which is an updated version of the software's architecture of the version currently implemented. In this diagram, the interaction can be seen between the system and the actors (stakeholders) involved in it and serves as a summary of the workflows of the whole process.

The solution developed has been in production since February 2014, monitoring continuously since that time (excluding time spent on upgrading the system to the latest versions) any relevant activity that has occurred in the laboratory. Correctly, the system has already monitored four real practices with students, in the courses 2013-2014, 2014-2015, 2015-2016, 2017-2018 of the subject Quality Assurance in the Laboratory Analysis in the Pharmaceutical Industry within the Master in Drugs Evaluation and Development of the University of Salamanca (Spain). The system has also monitored the interaction that other users have had during the time elapsed between such practical activities.

In such practice sessions, these tools presented here have been taken into account for the assessment of the subject, because teachers have decided to give a part of the evaluation of practice to the data provided by the system. Specifically, 25 percent of the practice mark directly depends on the percentage of critical points assessed by the student in the laboratory (% checkpoints achieved in Figure 11). The other 75% of the practice mark corresponds to the audit report presented by students at the end of their practice sessions, and where they show their knowledge about the GLP (Good Laboratory Practices) regulations and the elements of the laboratory that meet them or not. Other data taken into account when assessing the practical activities are a) if the student has carried out work on the agreed dates, b) if the student has made more than 1 hour of audit (in one or more sessions), c) if they have reached more than 12% of checkpoints revised. The teaching staff sets these three elements as the minimum requirements to show that students have audited and evaluated the laboratory.

Among the different results that can be shown to illustrate the operation of the entire system, it is possible to distinguish between two vectors of data to help test whether the system works, and to indicate if the system may be accepted for its actual use in teaching:

- Data concerning the validity of the system to help students practice and somehow replace a personal support teacher in the virtual lab.
- Data about measurements showing the usefulness of a system of collection, sorting, and processing of data in a case like this.

Table 15. Students and the average marks corresponding to the courses when practices were realized in the courses 2011-2012 and 2012-2013 (previous to the software) and 2013-2014 and 2014-2015 (using the software solution).

Course	Number of students	Average mark obtained (0 to 10)	Standard deviation
2011-2012	14	7.28	1.38
2012-2013	16	7.5	1.27
2013-2014	9	7.81	0.7
2014-2015	9	8.54	0.56

In the first case (*validity of the system to help students practice and somehow replace personal support teachers in the virtual lab*), to judge the results it is possible to compare the average marks obtained by students in the different years that the virtual laboratory has been used for practicing (years with direct support of the teachers as well as years with the implementation of this architecture and software tools). Table 15 shows a summary of the marks:

Based on these results, it is possible to see how students have increased their average mark (and decreasing the standard deviation) in each year. Are not shown no significant changes in the trend between the different years –namely, the disparities observed between those years where the practice was conducted as a single session with teachers' guidance and the other years where the practice was aided using the complementary technological tools developed within the scope of this project–. These data do not have to convey that the use of this architecture and the tools provided improve the results of practical sessions (this may be due to many factors, such as the decline in the number of students, the possible prior habituation of students to similar tools or 3D contexts, etc.). However, it illustrates how a trend of improving results is maintained even though the method of help, support and evaluation, of the practical activities.

As for the second information vector, about the usefulness of this type of system for monitoring, reporting results and aiding in the evaluation and development of practical work, a series of metrics can be displayed collected by the platform that illustrate the reader about the volume of information that the system has collected in these two academic years (2013-2014 and 2014-2015, specifically between 18 February 2014 and 20 March 2015), saving teachers from:

- 24276 interaction evidence (clicks in the laboratory).
- 249 sessions by 47 different users (teachers, students, and visitors to the facilities).
- 354269 seconds of interaction registered (total time of use of the system by the users, measured in seconds). It is more than 98 hours of interactions retrieved, stored and analyzed for use by teachers. Of this usage time, 137454 seconds (over 38 hours) of use of the laboratory correspond with students who carried out their practical work in the year 2013-2014 and 136581 seconds (almost 38 hours, again) of usage correspond to the time students who participated in the year 2014-2015 (in total, these students have registered almost 78% of total usage time in the laboratory).

These data are only a sample of what the Usalpharma architecture has contributed to the development of practice and the enormous task of knowing what happens in the virtual laboratory. These data are particularly relevant when analyzed in context: the data retrieved represent all the users' interactions over an extended period, regardless of the time zone of the visitor, the specific time of connection. The data provide metrics that teachers would never know or control without the help of the system.

Results of analyzing data and detect users' behavior and performance metrics

As mentioned the materials and methods section of this subchapter, the exploratory analysis of data retrieved from the Virtual World was performed using Microsoft Excel software. Once the proposed factors for engagement and usage were measured, they were counted, sorted, and clustered into groups of users, relating each user with performance and the engagement factor measures. These clusters and rules are obtained by counting and measuring the factors and joining primary factors with other related factors to obtain rules such as the following: *users who spend more time in the Virtual World use text-based chat more extensively*. To determine which groups or clusters of occurrences may constitute a behavior pattern, the following rule was used.

A behavior pattern can be established by observing significant overlap in the interaction and use measures by at least 15% of the users of the system. A significant overlap is established when a difference or deviation occurs in the measurements between users and other of at least 30% of the measurement.

This rule is based on considerations of association rules presented by Agrawal *et al.* in 1993 [238]. This research work, one of the main references in the field of association rules, explains the protocol for selecting relevant item sets and items within a dataset to create associations and rules regarding the data of a system. To select these relevant data, values and factors are used to discover rules; here, the author defined 15% of the population (minimum confidence of the association rule) as a substantial number of users with similarities in any measure and with a 30% difference from the same issue measured in remaining population (for example, a difference of 30% of the total movements between some users and the others); this 30% is the support of the rule and is based on the support difference between the minimum confidence (at least 15% of the population) of the rule and the other part of the population [239-241].

For cases with less disparity of factors and values (the presented case study has a small number of users with substantial disparity and heterogeneity in their characteristics based on different issues such as skills related to computer-interaction, age, etc.), values higher than 15 and 30% could be used to obtain higher accuracy in rule discovery, but in the present study, the author fixed these values to validate several rules to avoid empirical heterogeneity in the observations of the users' usage of the system.

To apply this rule, the author employed the following process with the software used to explore the data. Once the data retrieved from the Virtual World were entered in Microsoft Excel, it was classified in several columns according to the parameter measured (e.g., calls to API voice, instant messages, sessions, total time logged in, session number, average of time), and each measure was related to the corresponding user. Once the data were organized in the spreadsheet, the author searched for significant gaps between the data (Figure 12) such as users with a greater significant measure (30% gap) compared to the rest of the users (this gap facilitated the determination of the support of each rule that will be discovered later and must be larger than 30%). The author subsequently compared each value and selected users corresponding to coincidence in the significant use of each resource. When the author observed that more than 15% of the total population of the case study (15% was the minimum confidence of the future association rules proposed) exhibited coincidence in the range of the usage measure of the resources, a new pattern in the usage of the system was established, and a rule was formulated based on these data and the relationships discovered.

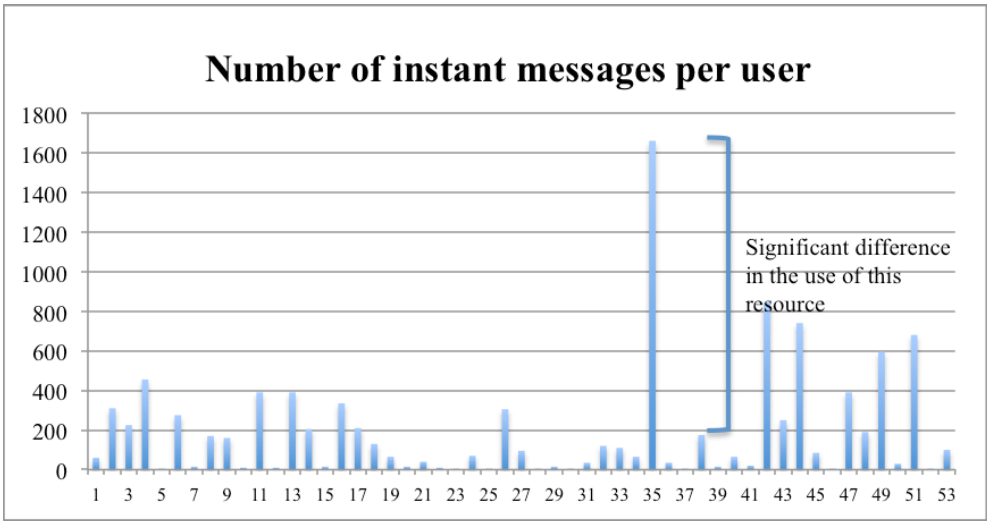


Figure 12. Example of significant differences in the use of a Virtual World resource.

For example, to consider that some users have a movement rate between islands of the Virtual World which may be considered as behavior pattern, this occurrence has to occur in at least 15% of them, and the difference in the measure of the movement (in and out of different regions) the selected group from the rest of users must be at least 30%. And if the pattern is expressed as a rule, for example, *users with more time spent in the Virtual World use more the text-based chat, both behaviors, more time spent and use more the text-based chat* must meet this rule of 15%-30%.

The exploratory analysis revealed many usage and behavior patterns. As explained, the primary patterns were then linked with related patterns (secondary patterns involving the same clusters of users). Additional patterns and characteristics that were identified in the data analysis but were not sufficiently representative or were only residually present were not included in this study.

After identifying the patterns, the author consulted through personal interviews with the expert panel to determine the accuracy of each feature or exciting pattern. The patterns and rules obtained from the data analysis and the opinion of the experts were as follows:

- *Users who participate in more sessions use more resources and tools in the Educational Virtual World.* Users who logged more time in the Virtual World (as measured by the number of logins and not time spent in the virtual world) often used a higher amount of resources (more interaction, voice and text messages, etc.). In particular, a relationship between the number of logins and text messages was noted. Interestingly, the number of resources used increases exponentially for users who connect often and is not a simple linear relationship; as the number of sessions and, possibly, acceptance of the 3D environment increases, users tend to perform a higher number of actions and exhibit higher ratios of actions per session. Of the 20 users with more sessions (in terms of total users and taking into account the 15%-30% rule), 12 exhibited this usage pattern (in Figure 13, each number on the X-axis is one user, and the Y-axis shows the number of sessions, voice API calls or instant messages, depending on the chart), corresponding to users 2, 3, 4, 16, 34, 55, 62, 63, 64, 67, 71 and 73. As shown in Figure 13, these users occupy the top positions in the ranking of use of text-based messages and voice features in the Virtual World.

The experts agreed that this pattern is a standard feature in Virtual Worlds as well as educational and other environments. The relationship between the usage and advantage of a tool can be expressed as a logarithmic curve, indicating that users who use a tool frequently know it profoundly and exploit it more. Related to the first heuristic rule, users from the third group of the rule (users who operate seamlessly in these environments), tend to make better use of the 3D environment and therefore spend more hours learning the features and options that the Virtual World offers.

- *Users who spend more time in the Virtual World tend to have more interactions with objects than with other users.* Among users who spent the most time (20 of the total users) in the Virtual World, more than half interacted more with objects inside the 3D scenarios than with other users of the experience.

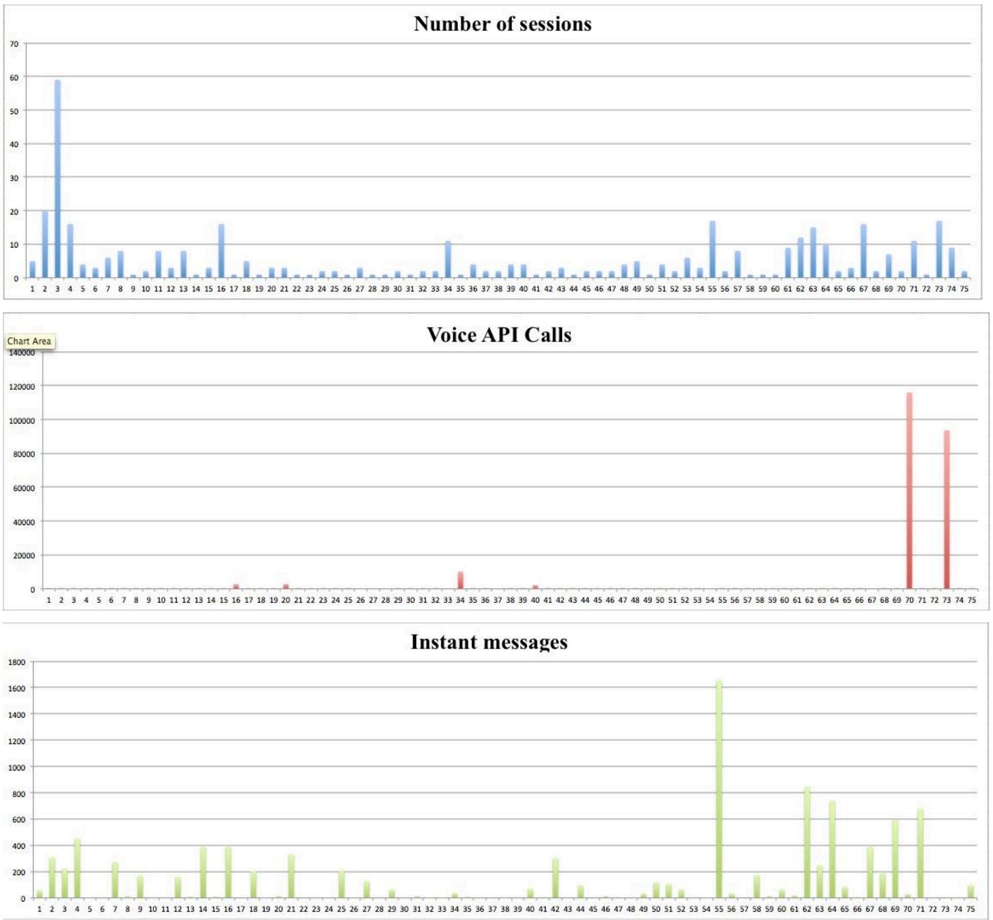


Figure 13. Different charts comparing the number of sessions, use of Voice API and number of Instant messages among users.

According to the experts, this pattern reflects the purely academic purpose of the Virtual World in which the data were collected. From their perspective, this pattern reveals that users who have spent more time in the Virtual World have spent more time exploring the 3D resources and solving the problems and learning challenges they entail. The experts assumed that users who use Virtual Worlds for hours are not always conducting practice or learning activities and thus often tend to explore the environments, usually interacting with the environment and the 3D objects. Without comparing it to academic performance (which could settle whether this engagement is based on academic aspects or similarity to video games), this pattern indicates that a user devotes time to the environment because he or she is interested in exploring it and improving the use of features included in the learning scenarios (second heuristic rule). The experts also indicated, as described in the fourth heuristic rule, that this goal is not always achieved because students must interact with other users in some learning activities (such as role-playing practices), and thus this result can be considered only as a behavior pattern in this Virtual World and not as a general rule applicable to any similar environment.

Teachers use a greater amount of voice resources. Six teachers were among the 20 users with the greatest use of the voice service within the Virtual World (the average number of Voice API calls by users and the standard deviations of these data are presented in Table 16). These six corresponded to users 3, 17, 20, 34, 70, 73. This statement is considered only a usage pattern and cannot be related to another pattern to establish a rule.

This pattern was confirmed in the fourth heuristic rule of the expert panel. Teachers use voice resources more because they tend to talk to students in these contexts, either to engage in conversation through their virtual alter ego (without the rigidity of the teacher-pupil relationship) or because many teachers are more comfortable clicking a button and talking than writing on the keyboard at the same speed as students (who usually write more quickly). For students, this difference in the use of verbal or written communication is often not obvious because they are much more accustomed to using text-based communication in instant messaging systems and existing social networks.

- *There is no relationship between the number of sessions and total time spent in the Virtual World.* None of the 20 users with the most significant number of sessions on the virtual platform were among the 20 users that spent the most time in it (Figure 14). Thus, many users connect several times for short durations, while others connect less often but for much longer durations; these two patterns are not directly related.

Table 16. Number of Voice API calls by the six teachers who were among the most active users of voice services within the Virtual World (average number of Voice API calls = 3085.7; Standard Deviation of the data related to Voice API calls = 16979.26).

User number	Voice API Calls
User 3	160
User 17	3513
User 20	2810
User 34	10190
User 70	115955
User 73	93625

The experts could not provide a clear explanation for this behavior. Two suggested that it might be related to testing experience with this Virtual World, such that many users are conducting connection tests in the virtual world, while the most skilled users performed fewer connection tests but remained connected for more extended periods. In summary, the experts did not perceive a regular pattern in these 3D environments, and the lack of relationship between the number of sessions and time may be either a temporary or unique feature of this environment.

- *Users who refuse this technology stop using it very early*. Of the 20 users who spent the least amount of time in this environment, 16 had spent less than 3 minutes online. In most cases, they had also used little or none of the other resources analyzed (voice, messages), as shown in Table 17.

All experts agreed that this type of platform produces very sharp reactions among users. As indicated in the first and third heuristic rules, users with no experience with video games or 3D environments frequently have experience difficulty in moving and using the avatar in the virtual environment, causing some to leave these systems very early. Moreover, the experts also suggested that students may not have taken the voluntary testing seriously and thus conducted only a connection test without actually experiencing right use of the platform.

Table 17. Comparison of total usage time (seconds), voice API calls and instant messages among the 15 users who spent the fewest time in the Virtual World.

Total usage time (seconds) Average (all users) = 1459.99 Standard deviation = 1758.9	Voice API calls Average (all users) = 3085.67 Standard deviation = 16979.26	Instant messages Average (all users) = 136.33 Standard deviation = 255.56
5	0	0
8	0	0
32	0	0
43	2810	15
49	0	0
68	70	15
96	60	0
115	80	0
119	80	205
125	40	0
126	50	60
131	90	20
144	160	225
156	0	390
161	30	210
168	40	160

In summary, and concerning engagement, the group of experts believed the most relevant patterns were the following: *users with more sessions use more resources and users who spend more time in the virtual world tend to interact more with objects than with other users.* These two patterns represent situations of pure engagement either in the platform or the learning content.

Besides, according to the panel, a greater in-depth analysis of object usage or voice-based chat usage might reveal if the engagement in these patterns is motivated by social relationships, the learning system, or purely entertainment based on its visual similarity to a video game.

For the pattern *users who refuse this technology stop using it very early*, there is no engagement with the system. This pattern could be handy for identifying reasons for lack of interest in the virtual environment among a group of relevant users and plan actions to avoid the dropout of these users.

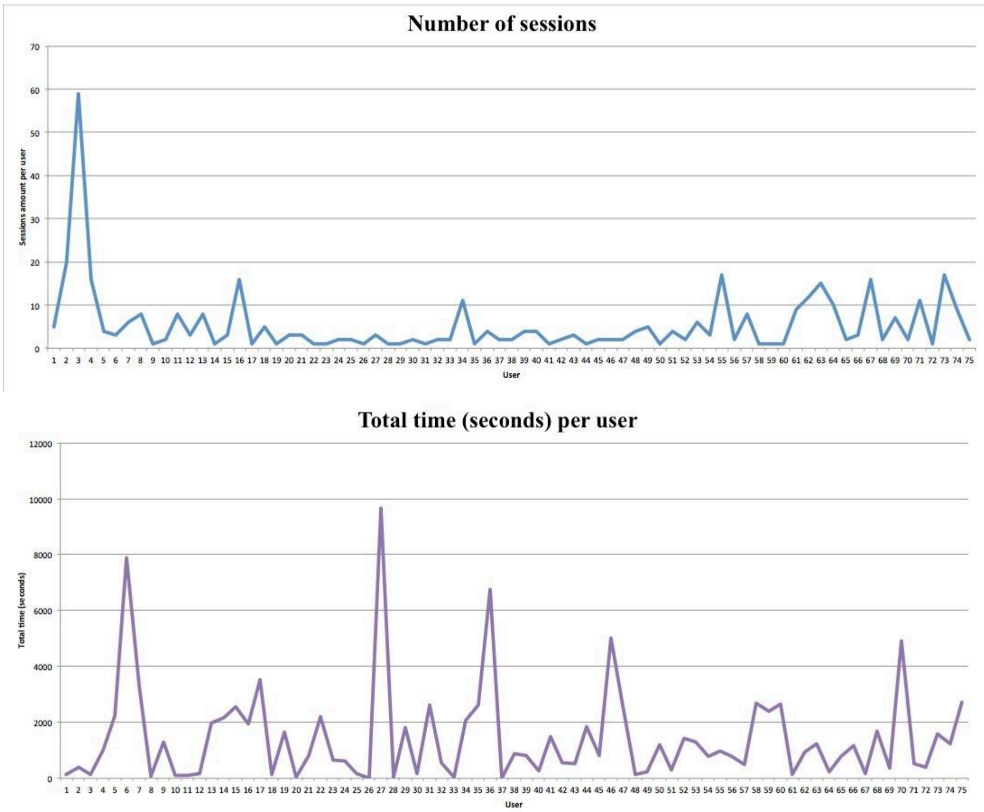


Figure 14. Comparison of the number of sessions and total usage time per user.

Moreover, experts related the pattern *Teachers use a higher amount of voice resources* with the teachers' ability to use the voice-based chat because it is simpler and faster for teachers to talk than write. The experts indicated that this usage of voice-based chat reflects the teachers' desire to communicate actively with students for, at least in part, academic reasons; such high use suggests that teachers are comfortable with the tool and have some engagement with it.

3.1.5 Discussion

About the research on gathering information and giving feedback to the users within a virtual world, first of all, it is raised an obvious question: is this proposal innovative? It is true that in the field of serious games there are already initiatives for the interconnection with systems like MOOCs [242], Personal Learning Environments (PLEs) [243], Virtual Learning Environments (VLEs) [244, 245], or Learning Analytics platforms [43, 246-250], but in the realm of virtual worlds, there are not many such initiatives [251]. A review of the current literature reveals that there is a tendency in the approach to proposals for the integration of different systems like VLEs, PLEs, architectures, and applications together with Virtual Worlds [207, 246, 252-255]. The literature also does not reveal the use in these cases of systems that analyze the performance of students within virtual practical activities or analyzing the interaction and the pieces of evidence collected from student usage of 3D learning scenarios and practices. The literature mainly describes the interconnection of Virtual Worlds environments like OpenSim or Second Life with popular platforms in the current teaching (Moodle, mainly) by using Sloodle and other platforms. These 3D environments extend their horizons of educational usage based on the support of systems and platforms where teachers already perform constant work creating materials and using them as an aid to teaching.

In fact, all the researchers involved in the experiments and experiences presented believe that this system is a real innovation in the field of teaching in virtual worlds (and even more in the field of education in Health Sciences). It is because of its features and vision outside the mere connection with systems that provide learning objects or materials created by teachers but as an aid to the evaluation and understanding of how students perform in such complex and dynamic environments like these [256]. Even in the most recent literature, it is possible to find authors (Griebel *et al.* [257], for example) that describe the proposal of this Usalpharma architecture as an exciting contribution that can open up new possibilities in the use of external systems completing the educational use of virtual worlds. These authors highlight that through the evidence collection system, the real-time responses both within the Virtual World and outside it, and its possibilities of use for the evaluation of activities. Therefore, it is possible to consider it a real, innovative and viable proposition for help in teaching. The interconnection of this architecture (using the services already available and others to be built) with other applications extends its field of education and eLearning and could help to achieve better results in learning scenarios arranged in 3D environments of this type.

The usefulness of such systems, for the author, depends primarily on the problem that it helps to solve. In the case of Usalpharma laboratory, this proposal provides information that cannot be achieved otherwise within a virtual world like Second Life, but it is necessary to consider whether this information is vital or not in the process of teaching and evaluation of students. In this case, the author agrees that, in the present case, the information is not vital, since evaluation depends mostly on the report submitted by the students (it represents 75% of the final mark); therefore, this system only provides extra assessment elements. What must be noted is that, in this particular case, those extra elements are also designed to establish a right to assessment or not (the reader will remember the conditions set for the assessment of the practical, of at least 12% of proven elements and a minimum time of interaction with the virtual laboratory of one hour). Moreover, the researchers involved also wish to state that the contribution made by the system in the form of information for students and teachers is a faithful overview of what happens in an environment like a virtual world, allowing teachers to measure aspects such as dedication (hours used), effort (number of sessions, number of actions to detect all elements that must be evaluated) or persistence (number of sessions used, evolution of the completion of the activity in each session, etc.), for example. This type of data, although not taken into account in the specific assessment for this practical work (and therefore the architecture is not decisive in the present evaluation) could be considered as catalyst values for the final evaluation. The data were part of the rubric for assessing practice work, giving specific metrics to aspects that are not typically measured transparently. In summary, although the system has not a vital utility for the practical work, it is increasingly being consolidated as a helper to the teacher. Additionally, the system is aligned with the current trends of tools, applications, and utilities that are opening new paths in new learning contexts that are relevant in technology-enhanced learning in the future.

Does the proposed system help to improve teaching? In this sense, it is difficult to provide a definite answer. By the data in Table 15, it is possible to say that the shift of paradigm and support in this case from classroom guidance to a system that provides feedback to students automatically, has not broken the trend of improved results, even when the results were improved significantly in the last year. The paradigm shift provided by this system is helping to improve teaching in a context such as a Virtual World, allowing students to experience an utterly immersive environment like Second Life and avoiding the guidance of the teacher in real time. However, venturing this conclusion would be going too far as it is possible that other factors affect results. Among the factors that could be relevant for the improvement of the results, can be highlighted the smaller number of participant students in the last years, the students' better predisposition towards such tools, the previous experience of some students with this kind of audit simulation, etc.

In any case, time will determine whether this system or others like this will significantly and unequivocally improve the students' learning and results, or automation of specific tasks previously performed by the teacher, are not really replaceable, and these systems directly must serve teachers as an aid in the acquisition of knowledge and decision making.

Can the system have more applications in addition to those described? This proposal is, even today, a proof of concept on the use and application of such systems and platforms related to eLearning. The author firmly believes that this proposal could have a more intensive usage and increased functionality through the integration with other systems and eLearning platforms, using more advanced algorithms and analysis techniques than those currently applied, etc. Especially using more advanced techniques, researchers could perform complex learning analysis, behavioral analysis (like in the other outcomes presented in this section), determination of the learning path for each student, personalizing learning, development of adaptive systems within Virtual Worlds, copy-detection among practices performed in 3D environments, etc.

About the research on discovering behaviors and engagement, first of all, should be asked the following: is this type of study useful? Does this study differ from similar studies of other systems used for learning? As many authors agree, among the community that engages in the formal use of Educational Virtual Worlds and retrieves information from these systems, many researchers and early adopters of this type of environment for education have recognized a gap between the events inside the virtual world and the teachers' knowledge of these events [171, 192, 226]. This gap cannot be addressed easily. The data and evidence retrieved from other types of environments are very different from those that can be collected from Virtual Worlds; these differences may be due to the nature of the student fingerprint, which requires new approaches to understand user interactions, or to the private nature of some of these environments [91, 170, 171, 224]. This type of study is useful because obtaining knowledge about the users' usage and interaction with the Virtual World and learning through the contents and resources presented within a Virtual World helps teachers and managers of the Educational Virtual Worlds to better plan and design the deployment of educational content and resources inside the 3D environment, enhancing the personal experience and learning process for students.

Accordingly, how can managers use this knowledge to improve the user engagement and the utility of the Virtual World? Based on the cases and rules described above, there are several rules (those based on the usage behavior) that managers of the virtual environment could follow to help users make better use of the environment and meet their learning goals.

For example, in the case of the rule *Users who refuse this technology stop using it very early*, managers could plan actions, reminders, classroom activities to help students, and, if this proportion of students is considerable, plan introductory workshops about this technology. Knowledge about early dropout from this technology and environments allows managers to remain vigilant about users' engagement indicators and act to avoid dropout among users meeting this rule.

In the case of the rule *Teachers use a greater amount of voice resources,* the managers can focus their attention on voice chat usage indicators to identify teachers who are not using voice chat and may be experiencing technical problems or are unsure how to use it correctly.

In the case of rules reflecting engagement (for example, *Users with more sessions use more resources; Users who spend more time in the virtual world tend to interact more with objects than with other users*), these indicators could serve to measure interest in the different features available in the 3D environment to allow managers, teachers, designers and developers of the virtual world to successfully design new learning scenarios that are more interesting to students or the general users of the Educational Virtual World. These indicators of engagement and types of user profiles can also allow managers, designers, etc. to design activities specifically for each type of user. Thus, different features and scenarios could be offered to users who use and perceive the environment differently to facilitate learning goal achievement.

3.2 Analyzing MOOCs users' interaction to get insights about their learning processes

The emergence of the Internet and the eLearning concept have modified the way humans learn and interact with knowledge [244, 258-261]. Specifically, this change has undergone a remarkable process of acceleration with the emergence of new theories, methodologies, tools, and systems. These new proposals, designed and implemented to take more and better advantage of the online medium, facilitate the acquisition of knowledge by students and reinforce their learning processes, independently of age, sex or other personal conditions [262, 263]. The evolution of technological ecosystems [264] intended to support eLearning has given rise to environments capable of adapting to the users [265-267], integrating gamification techniques [268-270] to the process of acquiring knowledge, offering collaborative learning tools [271], analyzing their learning and acting accordingly to reinforce it [256, 272], or providing the ways to greatly increase the number of students who can study a subject simultaneously on a virtual platform [273]. Especially relevant are these latter learning environments, better known as MOOC [274]. These learning environments make available to many users (thousands in many cases) virtual classes as knowledge containers that provide open learning resources for all users enrolled in the course, as well as methods and systems to reinforce that knowledge acquisition from different perspectives [275-277].

Furthermore, in addition to the revolution that has introduced eLearning in the field of learning, it is necessary to emphasize that the boom in systems and methodologies such as those used to support the MOOCs is due in large part to the fact that learning is not currently conceived from the classical formal point of view. According to the literature [278-285], at present, three types of learning can be distinguished:

- Formal learning is one in which learning occurs in organized and structured environments (in an educational institution or a work environment) and is explicitly designed as learning concerning objectives, time and resources. Typically leads to processes of validation and certification on the knowledge acquired.

- Non-formal learning is one that is embedded in planned activities without being explicitly posed as learning. Nonetheless, non-formal learning can occasionally be validated and lead to certifications.
- Informal learning is one that results from daily activities related to work, family, or leisure. It is unorganized and unstructured learning regarding objectives, timing or support of conventional structures related to learning.

The combination of these types of learning (especially non-formal and informal) with systems and methodologies purely related to eLearning have opened, and continually open today, new perspectives on the training of the individual of the 21st century from a continuous point of view in time, multivariate in terms of resources that can be consumed and ubiquitous in terms of availability. At present, when information overflows the individual, having the vehicles to transfer real and valuable knowledge to the people can be one of the cornerstones in building a more rational and advanced society.

Among these information vehicles, many hypermedia resources have appeared over the years. Among them, the most important are websites, forums, blogs, and the social networks [286]. These social networks support true learning communities [287] where conversations are developed; content is shared openly, relationships are established among users (in a horizontal way in many cases), interaction among people and digital entities appears, etc. In other words, social networks constitute a real Petri dish where users form digital societies by developing communication structures, consumption patterns (in this case, of information) and user networks in a broad way. Is in these social networks, in these digital societies, where some of the aspects mentioned previously can come together: eLearning in a non-formal or informal context, as well as formal [288]. As discussed by various authors [113, 289, 290], informal conversations and content present in social networks are currently one of the most successful sources to acquire extra knowledge and improve the learning experience in online courses. Furthermore, individual conversations and interactions that take place in social networks can be a consequence of the realization by social network users of some learning activity (online or offline) [291]. In this sense, it is possible to refer to the Connectivism theory [292, 293], which argues that the learning process is enriched by the connection of students, teachers and online resources. Additionally, social networks are a perfect means to improve this connection [294, 295], thus favoring the emergence of true communities of connected learning and practice [296].

In the case of the research presented in this section, in the same way as in the rest of experiments, most parts of the results have been published. These publications [43, 250, 297] are available in the Appendixes 6.5, 6.6, and 6.7.

3.2.1 Context

The analysis of interaction among users and systems provide significant insights into how users use, understand and take advantage of tools and platforms they utilize to perform any task.

The fact of analyze the interaction and try to extract valuable knowledge from it, have real application in many areas of knowledge and business, as in digital marketing, in education (Learning Analytics), etc. Although in some fields as education, this type of behavior analysis, and interaction analysis is increasingly common, the approaches and tools developed should be updated and adapted to new systems, platforms, and paradigms in eLearning. In these new types of learning platforms and paradigms can be highlighted the MOOCs, because they expand the traditional limits in students' interaction with teachers, contents and online learning platforms. Furthermore, MOOCs leverage other platforms (even those that are not purely intended to be applied in education) like the social networks and other online tools, applying by this way multi-platform and multi-context approaches that can improve and upgrade the learning experience [298].

It is because of this use of multiple tools and multiple contexts that is necessary to design and implement new ways of interaction analysis and platforms that allow to perform it. These new ways and platforms will manage the acquisition of knowledge regarding the learning and interaction with platforms, establishing convergence of knowledge between different learning vectors and context, to finally allow teachers and managers to learn, explore and implement possible improvements that help in the learning process, the design of content and the motivation of students.

3.2.2 Goals

As presented in section 1.2, the objectives pursued in this scenario are:

1. Understand how users collaborate in a large environment to solve (individually) a task shared by all users.
2. Integrate a software solution in a massive environment to automatize the analysis performed in the primary objective.

These goals can be specialized in the case of this experience with MOOCs to:

1. Design and develop a modular software architecture to allow teachers and managers of a MOOC retrieve knowledge about how users enrolled in a MOOC course utilize some tools external to the MOOC platform. By this way, these stakeholders could get insights about what did users on these external tools, what kind of interaction they perform inside them, and thus, discover possible improvements and solutions for eLearning processes to be applied later in the MOOC platform and its courses.
2. Discover how is performed the knowledge acquisition by users when interacting with other peers in environments non-designed for Learning like the social networks [287].
3. Study users' conversations and interactions with MOOC content or related on social networks like Google+ or Twitter and how users use tagging resources on such networks (hashtags) [250].
4. Determine if there are patterns or coincidences between the use of social networks and the advance in the MOOC courses by the users.
5. Evaluate if it is possible to use the users' usage of MOOCs and social networks to establish parallelism between both and to determine the types of learning that are given in these environments.

3.2.3 Materials and methods

The MOOC platform that has been used to carry out the study is iMOOC [299, 300]: a system developed by the Polytechnic University of Madrid, the University of Zaragoza and the University of Salamanca. iMOOC is based on non-formal and informal learning and has characteristics of adaptability, gamification, or collaborative learning among others.

The iMOOC or intelligent-MOOC platform includes, among other actions, the creation of a MOOC platform based on adaptive learning and information [281, 301-303]. To achieve this goal, the project uses the eLearning Moodle platform (https://goo.gl/py8cJa), specifically version 2.6.5, taking advantage of its great versatility.

This adapted learning is possible thanks to the use of the different tools offered by the platform. These tools include conditionals, groups and user groups that allow creating and then choosing different groups by associating them with the different course resources. The use of these strategies and tools can lead to different educational itineraries depending on the type of profile of the user, path chosen according to the theme or the progress of the student within the course and level of knowledge.

On January 12, 2015, a first demo of the iMOOC platform was launched with the course "Social Networks and Teaching", a special version of the course "Application of social networks to teaching" previously developed in the platform MiriadaX (https://goo.gl/EG1sEF).

This course is based on the cooperative model defined by Fidalgo *et al.* [304], which contains characteristics of the two more standardized types of MOOCs such as xMOOC, with a more behavioral and similar approach to traditional online courses compared to cMOOCs whose approach is more connectivist [305-307] based on the social networks. To explain this cooperative model, the course can be divided into a series of layers, starting with the "technological" layer. This layer includes, on the one hand, the MOOC platform that hosts the course and on the other, the social platforms where the interactions between participants and the content generated by these two. This layer is followed by the "formative strategy" associated with the instructional design of the course and finally the "cooperative" layer that represents the most connectivistic part of the course, gathering the results and the content generated from the cooperation between the teaching team and the participants in the course and integrating it with this. It is necessary to add to the defined cooperative model the fourth layer to explain the gcMOOC cooperative-gamified model on which the course is based; this layer is called gamification [308], which interacts with the other layers, fostering the motivation of the participants in the course [308].

As for the content of the course, it aims to introduce the student to the social web for a month, identifying the phenomenon of social networks within it and the opportunities they offer within the field of education, more specifically, in the generation of virtual learning communities. Otherwise, the use of the most widespread social networks such as Facebook, Twitter, and Google+ is explored for the student to develop the digital skills necessary to deal with this type of technology, while offering a series of guidelines for application in the classroom. Finally, it gives an overview of a total of 13 other social networks also suitable for this teaching use, as well as tools for more optimal management.

Based on the characteristics offered by the iMOOC platform, students are offered five different itineraries that they can choose before starting the course, adapting it to their needs:

- Full course for teachers (with two additional lessons focused on the use of Twitter and Facebook as an educational tool).
- Full course for non-teachers (without teaching lessons).
- Twitter course, formed only by the module of the said social network.
- Facebook course, a single module on this network.
- A special itinerary focused on those students who participated in any of the previous editions of the course, with an additional module focused on the practical use of the learning community. This itinerary also offered the possibility of repeating the complete course, allowing its students to access all the contents.

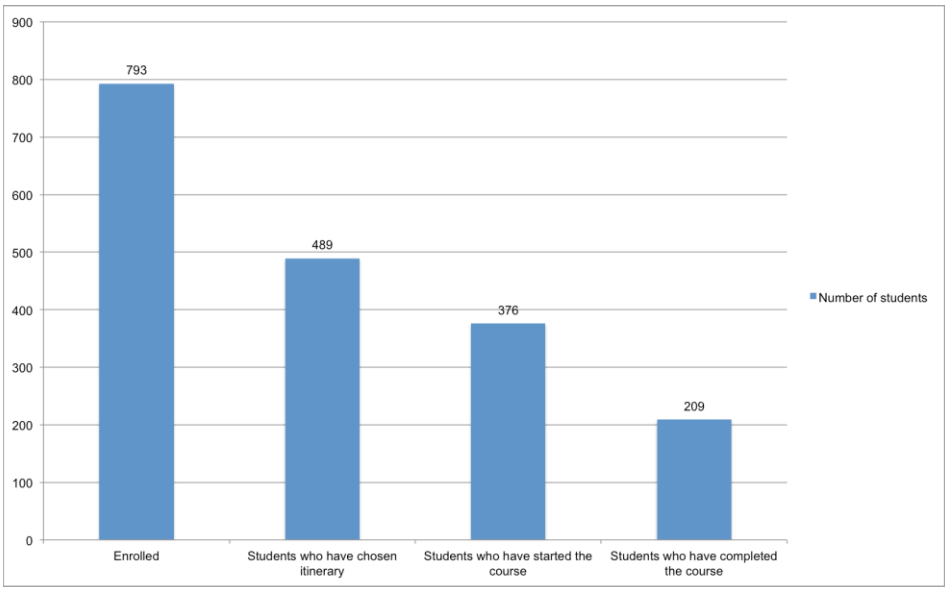


Figure 15. Distribution of students regarding enrollment, itinerary selected, initiation and completion of the MOOC course.

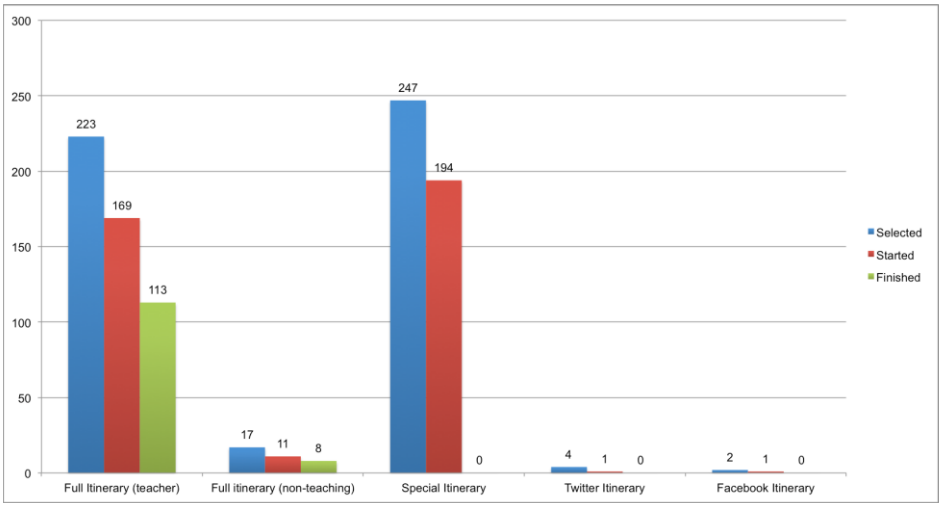


Figure 16. Distribution of students regarding the learning itineraries available in the course.

As general data of the course follow-up and some metrics about the itineraries and completion of the course, the following summary is offered to the reader (Figure 15, Figure 16, Figure 17). In the specific case of Figure 17, those students who chose the special route to obtain the certificate were excluded from the metrics. It is because they had only to present the certificate obtained in either of the previous editions of the course in MiriadaX, to visualize the contents of its special module, based on an implementation of learning communities from two different points of view and participate in the forum. Even so, the students of the special itinerary were offered the possibility of completing the full MOOC. Of the 188 who made this itinerary, 107 repeated the course, and of them, 88 passed it (82.2%).

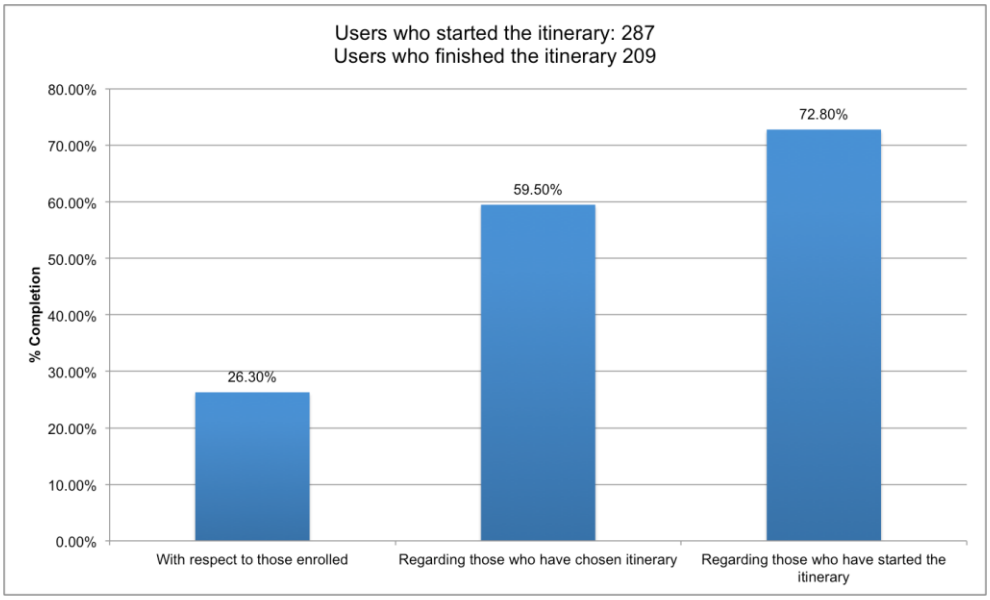


Figure 17. Statistics on the initiation of completion of the MOOC course without taking into account the special itinerary, because this itinerary did not require the same effort as the rest.

In the case of a course that deals with "social networks in teaching", and given the connectivist approach of the MOOC and the course, the social networks have played an essential role in the learning process associated to the course. These networks have been used from two different perspectives. The first is the use of social networks as test environments and as case studies to obtain a practical understanding of the concepts theoretically shown in the course. The second perspective is its use as a platform to continue and extend the learning process of the iMOOC course from both a non-formal and informal perspective. To accomplish this second perspective, teachers proposed suitably labeled conversation topics (through hashtags), opening up ways of discussion and acquisition of knowledge from a non-formal point of view, encouraging the emergence of conversations and informal learning. The students proposed the communication channels in an environment other than the MOOC, as well as for discussion with other social network users who participate in the conversation without being enrolled in the course [309] as shown in Figure 18.

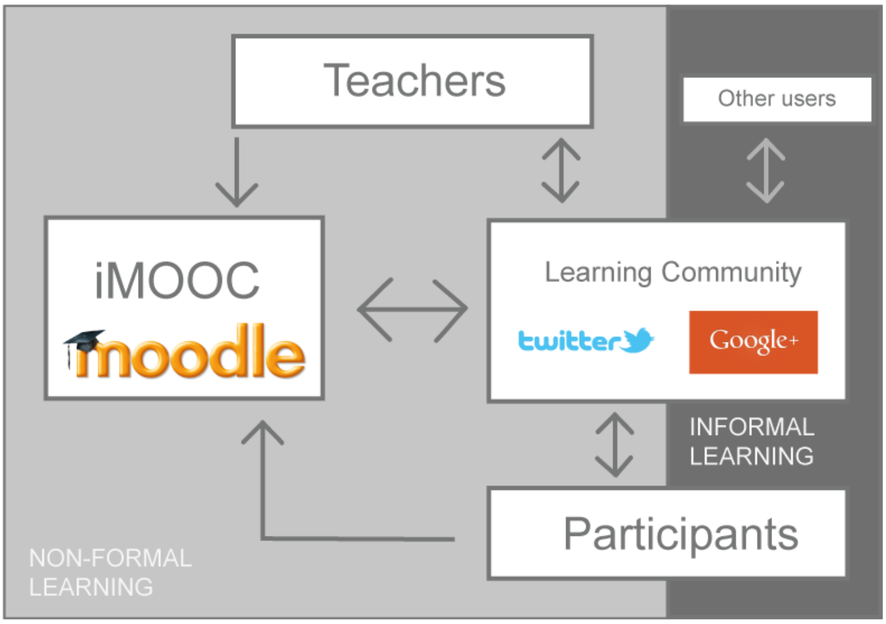


Figure 18. Schema on the interaction between users, iMOOC and social networks segmenting learning according to the type that occurs in each phase.

For this task, this MOOC has used social networks Google+ and Twitter. In Google+ this course has been associated, since its inception in the MiriadaX platform, to a community with more than 5000 users (https://goo.gl/1SZMFD). In this community, users share resources and collaborate with them in the learning associated with each edition of the course. Regarding Twitter, it has been used as an alternative social network to raise discussions and conversations between users, due to its current popularity and the facilities it offers regarding tag, track and information retrieval from users' conversations [43, 310, 311].

To obtain information about the interaction of iMOOC users with the contents of the course in social networks, it is necessary to establish adequate information retrieval channels [42]. Specifically, in this case about the relationship between iMOOC user profiles and different social networks, as well such as retrieving the information they share and the tagging, they perform on those contents, to be able to carry out the analyzes that are presented as study objectives.

As a summary of the methods used, can be highlighted:

- To avoid manual retrieval of each user's data, web services have been used (Representational State Transfer —REST— [312] APIs) that offers the Moodle platform on which the iMOOC is based. Using the REST API, user registries and their profiles have been accessed, allowing the author to filter those who have registered their profiles in social networks (condition proposed by the MOOC teachers to help in the evaluation of students) [43].
- To extract information from Twitter, the automatic recovery of tweets (through its REST API https://goo.gl/2aG5sN) has been combined with the manual recovery of some specific metrics.
- As for extracting information from the social network Google+, due to the lack of APIs to retrieve information from user communities (https://goo.gl/sYZ2Ma), it has been necessary to develop a tool called GILCA (Google Analytics Informal Learning Communities). This application collects data from Google+ communities through the email notifications sent by the social network (including information on publications, comments, hashtags, etc.) [43].
- To further understand how users use the tagging functions in the social networks used in this article, they conducted a questionnaire that addressed fundamental questions about the use of hashtags in social networks and activities.

To develop the software architecture, the technologies used were:

- Django Framework [231]: This web framework is used to build the software layers and to coordinate the information workflows between the components and systems of the architecture.
- MongoDB [233]: This NoSQL database is used to store the data without the traditional restrictions of the SQL databases, and allows to adapt the database storage schemas to each kind of content retrieved from external tools and platforms [177].

- REST APIs: these web services are used to serve as communication channels between components and systems involved in the MOOC ecosystem. In case of those tools and systems that do not provide this kind of facilities, are used crawlers to retrieve information (this is explained in the following section).

As for the analysis of the data obtained, this was done with spreadsheet tools.

3.2.4 Results

Results related to the design and development of a software architecture to extend data-driven MOOC technical features

Following the previous experience of the author in similar cases, where he applied software architectures to extend the functionality of eLearning ecosystems [176, 177, 226, 234, 256], the author decided to use the core of a software architecture they built in 2014. Several layers compose this core, one to retrieve data from each external platform or tool, other wiping and stores the information retrieved, another to push analyzed information to other platforms, and others that enable searches and interaction between information and users.

The core of the architecture is a system that acts as a mediator between the different social networks and learning platforms that are interconnected (Figure 19). This mediator system communicates with each external tool through using web services (REST APIs commonly) and crawlers; retrieving data and information from them and analyzing the information to convert the raw data in valuable information for teachers and managers of the iMOOC platform (based on Moodle).

The main idea behind the framework is that teachers and managers could use the website provided by the architecture to interact with the information and data retrieved from the external tools, so all the assessment and evaluation of the users learning in the MOOC could be centralized in the architecture. The other possible approach is that the architecture pushes the analyzed information again to the Moodle platform, allowing teachers and managers to allow them to assess and evaluate the learning without leaving the MOOC platform, in that case, the data visualization and interaction with the information retrieved depends strictly on the resources Moodle provides.

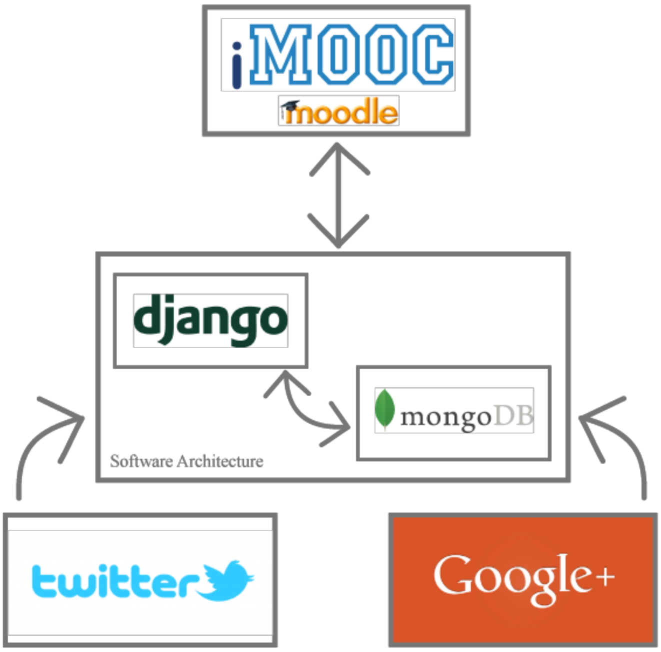


Figure 19. Software architecture proposal for the experiment related to MOOCs.

To implement the information workflows shown in Figure 19 between the software architecture and the different systems and social networks, researchers need to establish the proper communication channels for the information. These communication channels are based, in this case, in services and crawlers:

- The services are facilities provided by the third-party software to facilitate the communication and interconnection with other systems, applications or clients. In this case, researchers have used services for communicating with Moodle and Twitter.
- The crawlers are software applications that find information automatically in third-party systems when they do not provide services for the pull and push information between systems. In this case, the author was working on crawlers for getting information from Google+ Communities (Google+ does not provide API or other services to get and post activities and other information within the communities).

Below is explained how have been used these services and crawlers, and how they are implemented within the software architecture.

Moodle provides several API services and API architectures; allowing users and third-party applications and systems interact with courses, administration settings, users, and configuration information. The API used in this work is based on a REST architecture, and it allows several actions in both directions of communication (GET and POST actions, as well as DELETEs, etc.); the full documentation and functionalities of this API can be found at [313].

For example, these API endpoints and functionality allow managers and teachers of the (i)MOOC course to make automatic checks about the tasks completed by users, automatized (and directly) assessment about their participation in the MOOC, etc. In a regular course on Moodle, this usage of the API is not a critical aspect, most of these checks and assessment are performed manually by the teachers. However, in a MOOC course with more than 700 hundred (in this case, several thousand in bigger MOOC courses) turns out this resources as a critical factor to evaluate the users' interaction with the MOOC and for assessing their learning within the course.

Below is presented an example code that allows teachers to retrieve the full list of users enrolled in the course. This result list of users enrolled, for example in the case of the iMOOC course was used to check what users filled their profile with the links to their personal social networks profiles, which was proposed as an activity of the *Twitter in education lesson*. As previously explained, author implemented the software architecture using Python language, so the code is formatted in a *pythonic* way and includes the main software library used, Requests [314] that allows implementing the API consumption efficiently.

```
import requests, json

parameters = {'wsfunction': core_enrol_get_enrolled_users',
            'courseid':'id',
            'moodlewsrestformat':'json',
            'wstoken':'xxxxxx'}
url = "http://gridlab.upm.es/imooc/"

response = requests.get(url, params=parameters)

if response.status_code == 200:
   results = response.json()
   for result in results:
      print result
else:
```

```
print "Error code %s" % response.status_code
```

Regarding the Twitter data retrieval implementation, the author has implemented collector that gets tweets on real-time based on their hashtags. This implementation is possible thanks to the Twitter REST APIs [315] and Tweepy library for Python [316]. Using both facilities (especially the Twitter Streaming API), the author built software that can retrieve in real time tweets tagged [310] with the any of hashtags proposed in the MOOC course and storing the tweets in the software architecture database (enabling by this way MOOC user matching, etc.). An example of how is done this data retrieval, below is attached a simplified version of the code:

```
from __future__ import absolute_import, print_function
from tweepy.streaming import StreamListener
from tweepy import OAuthHandler
from tweepy import Stream

consumer_key="xxxx"
consumer_secret="xxxx"
access_token="xxxxx"
access_token_secret="xxxxx"

class StdOutListener(StreamListener):
    def on_data(self, data):
        try:
            print(data)
            return True
        except:
            pass
    def on_timeout(self):
        sys.stderr.write("Timeout, sleeping for 60 seconds...\n")
        time.sleep(60)
        return

if __name__ == '__main__':
    l = StdOutListener()
    auth = OAuthHandler(consumer_key, consumer_secret)
    auth.set_access_token(access_token, access_token_secret)
    stream = Stream(auth, l)
        stream.filter(track=['#RSEEjemplosRRSS',
                    '#UsosTwitterEnseñanza',
                    '#RSEMiKlout'])
```

About Twitter integration in the system, should be highlighted that the MOOC managers and teachers must get the permission of the users about storage their tweets, or merely anonymize the personal data present in each tweet (name and username, location, etc.) because the social network specifies in their policy rules this restriction.

About Google+, the situation is entirely different. This social network provides APIs and methods for retrieving information about users, posts, comments, etc. [317], but it does not allow to retrieve information from the users' communities within the social network. This disables the possibility of using the same way to get information about conversation and interactions in the communities, regarding this, teachers and managers from the MOOC course were searching other tools that let them retrieve the desired information; for example, they use currently the tool AllMyPlus (https://goo.gl/jfWizP) that allows them to retrieve information of the learners community related to MOOC. This is not the best solution, because it converts the ideal automatic process indeed in a manual process, so the author was trying to develop a crawler that enables them to retrieve information directly from Google+ website or AllMyPlus website.

Results related to the analysis of users' interactions and detection of learning communities in the MOOC and social network related

As mentioned, the questionnaire filled out by course users contained questions related to the use of labels on social networks [318]. 212 users, 26.73% of the total enrolled in the course, completed this questionnaire. In the following figures (Figure 20, Figure 21, Figure 22, Figure 23, Figure 24, Figure 25) can be seen the data retrieved through these questionnaires and the segmentation of responses according to the age of the users.

To obtain results from the use of social networks, researchers filtered what users had indicated their Google+ profile or Twitter in their iMOOC profile. This filter allowed to retrieve which of them had been published in social networks following the official hashtags of the course (Figure 26). Once this check was carried out, the researchers evaluated the number of publications they had made and the proportion of users published on Google+ or Facebook. Finally was evaluated if these users with social interactions and social profiles had approved the course or not.

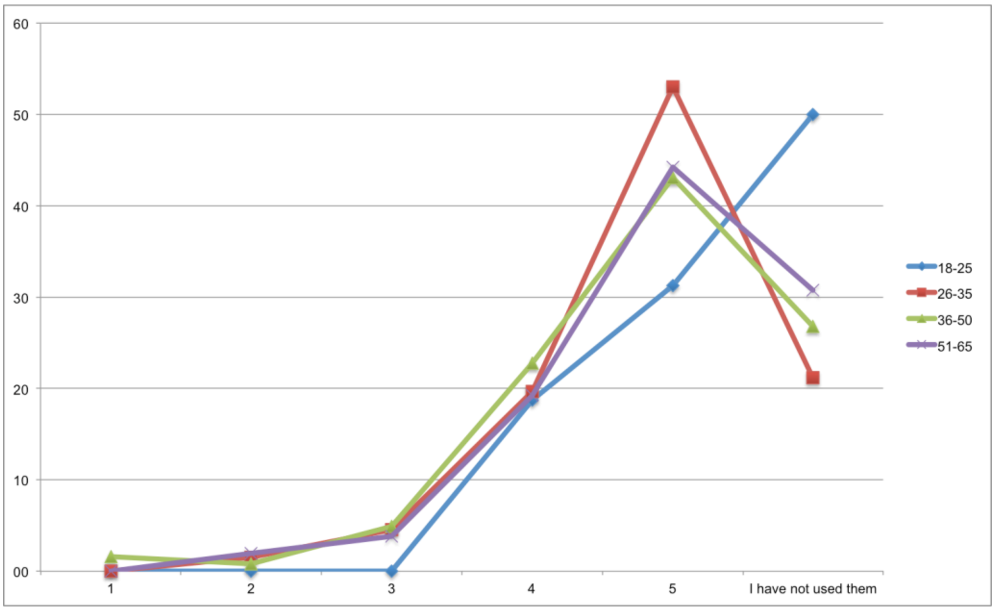


Figure 20. Results of question 1 of the iMOOC questionnaire: Evaluates the utility of using hashtags (Likert scale 1-5 plus null value). Age-targeted responses of users.

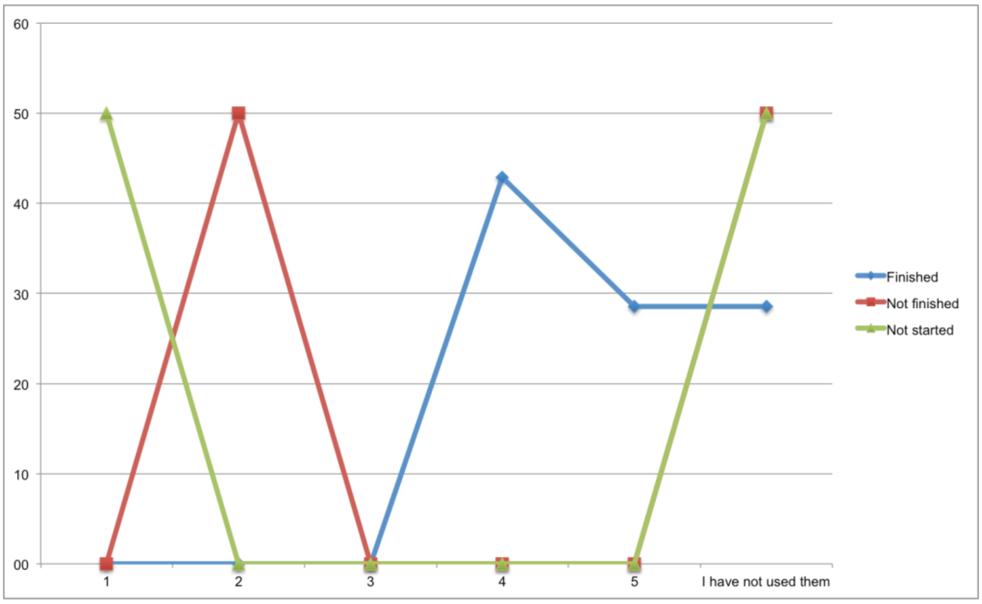


Figure 21. Results of question 1 of the iMOOC questionnaire: Evaluates the utility of using hashtags (Likert scale 1-5 plus null value). Responses segmented by degree of completion of the MOOC by the users.

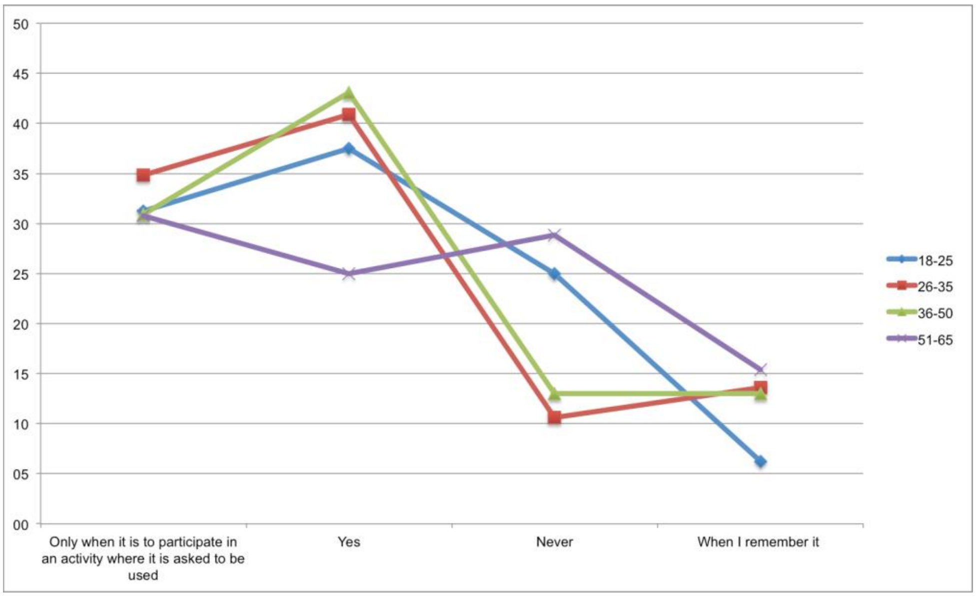


Figure 22. Results from question 2 of the iMOOC questionnaire: Do you often use hashtags in your publications? Answers segmented by age groups of users.

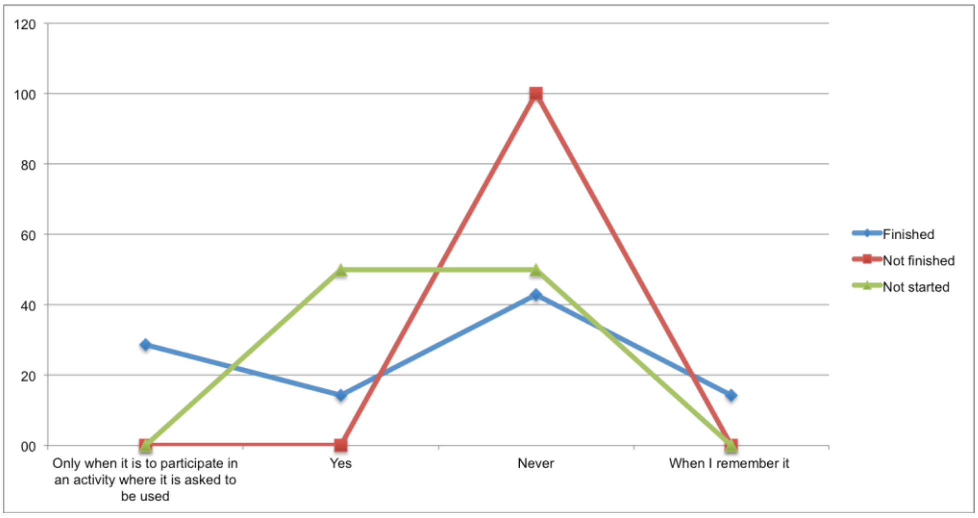


Figure 23. Results from question 2 of the iMOOC questionnaire: Do you often use hashtags in your publications? Responses segmented by degree of completion of the MOOC by the users.

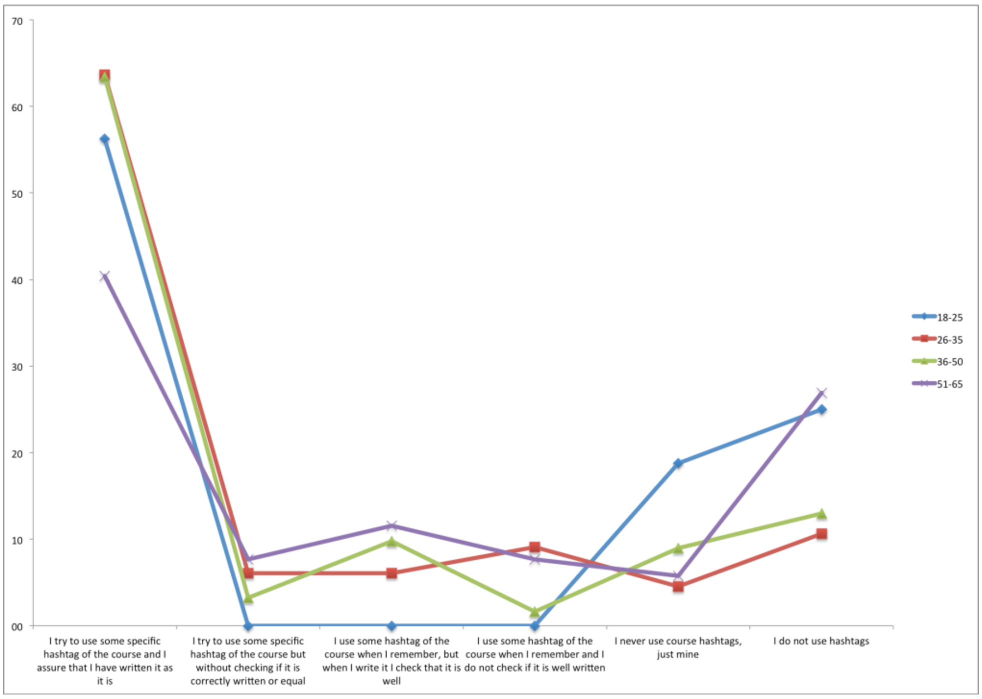


Figure 24. Results from question 3 of the iMOOC questionnaire: "When publishing and using a hashtag ...". Answers segmented by age groups of users.

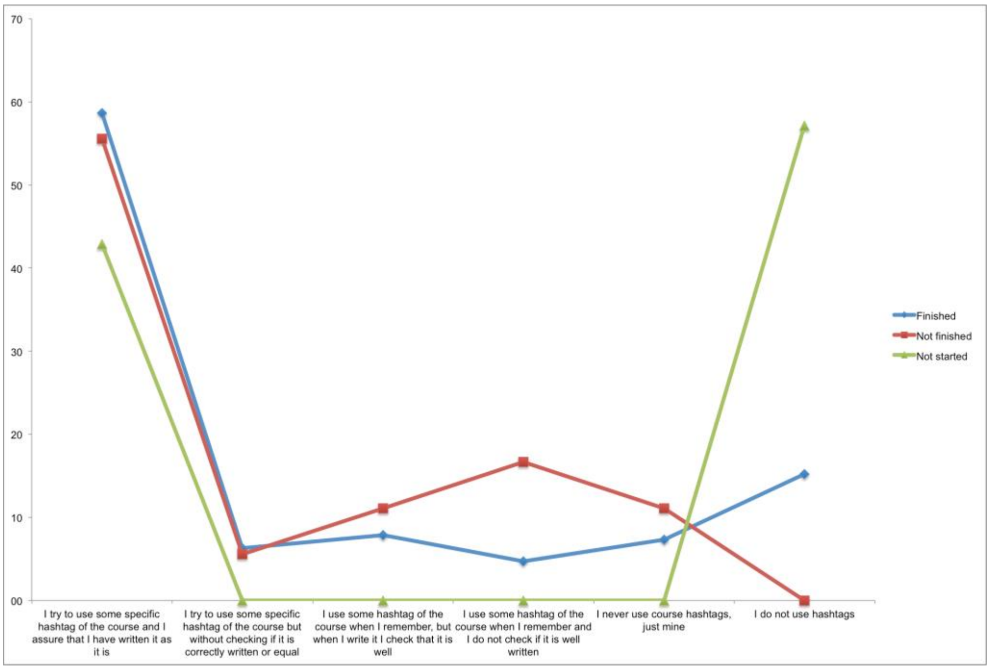


Figure 25. Results from question 3 of the iMOOC questionnaire: "When publishing and using a hashtag...". Responses segmented by degree of completion of the MOOC by the users.

This evaluation resulted in (Table 18):

- The students enrolled in the course have made a total of 263 publications in the Google+ community, also commenting on other peer publications.
- The students enrolled in the course have published a total of 131 tweets following official and unofficial hashtags on Twitter.
- Of the users who have made a publication in Google+ (191 users) have passed the course 57 (29.84%).
- Of the users who have published a tweet on Twitter (76), have passed 42 (55.26%) the course.
- Of the 191 users who indicated their Google+ profile on iMOOC, 83 users (43.5%) have passed the course.
- Of the 265 users who have indicated their Twitter profile on iMOOC, 105 users (39.62%) have passed the course.

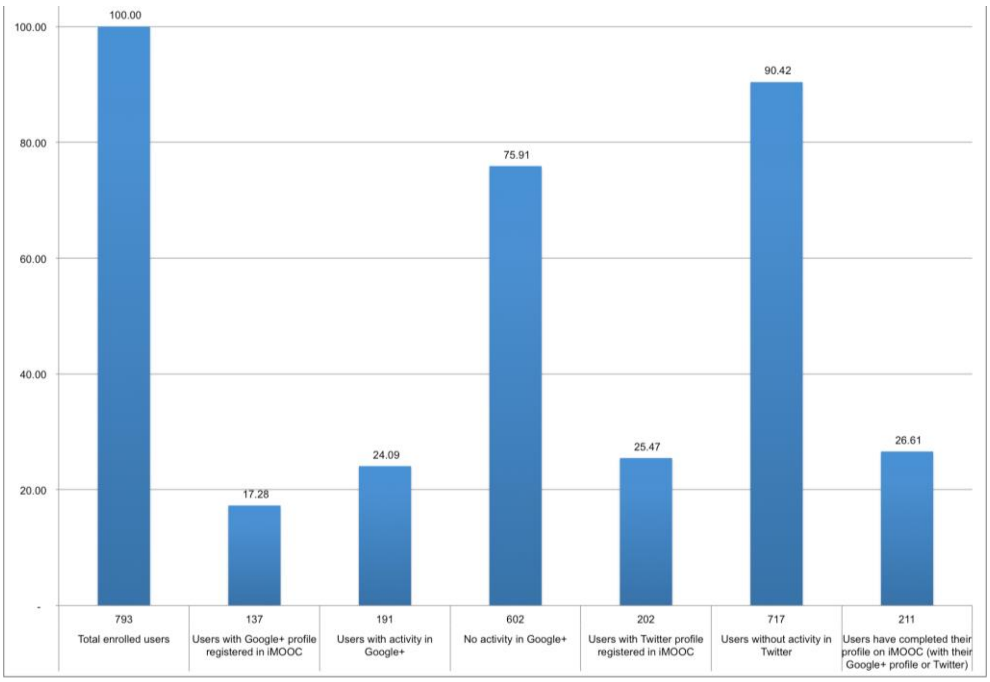


Figure 26. Distribution of iMOOC users regarding their use in social networks.

Table 18. Distribution of interactions on Google+ and Twitter by type of content and learning.

Google +						
Publication						
Category	Type	Amount	+1s	Comments	*Reshares*	Type of learning
Debates	Proposed throughout the course	1	83	17	14	Non formal (proposed by teachers)
	Use of social networks	4				Informal (proposed by students)
	About learning	3				
	About digital identity	2				
	About digital identity	1				
	About Facebook	1				
	About *badges*	1				
Total posts in the discussion category = 11						
Activities and exercises	Examples of social networks	31	309	41	20	Non formal (proposed by teachers)
	Exercises on bad practices in social networks	25				
	Exercises about Facebook	28				
	About influence (Klout)	22				
	Uses of Twitter in teaching	3				
	Others	2				
Total publications on activities and exercises = 111						
Resources	150		552	66	93	Informal (proposed by the students)
Twitter						
Publication						
Type	Hashtag	Tweets	Responses	Retweets	Favorites	Type of Learning
General	#RSEMOOC	9	2	5	5	Non formal (proposed by teachers)
	#RSEHANGOUT	19	4	16	15	
	#Modulo1RSE	1	0	1	1	
	#Modulo2RSE	1	0	1	1	
	#Modulo3RSE	1	0	1	1	
Activities and exercises	#RSEEjemplosRRSS	4	1	0	0	
	#RSEMalasPracticas	5	0	1	2	
	#RSEmiKlout	8	1	5	6	
	#RSEMoodleTwitter	59	9	9	11	
	#ActividadesRSE	1	0	3	3	
Total tweets = 107						

3.2.5 Discussion

Both hashtags on Twitter or Google+ and the categories in the Google+ network represent an opportunity for collaborative MOOCs. Also, they are an opportunity to those MOOC with characteristics related to non-formal and informal learning, since they allow to classify and collect the contents generated in the communities related to the course, and to feedback the MOOC from this non-formal and informal content. The main drawback of retrieving this labeled knowledge, as can be observed in the results, is the lack of digital skills, custom, and awareness on the part of the participants in this learning communities. Depending on what is observed, usually, the result of the conversations are orphaned publications regarding labels and even in many cases finding that these labels do not match the ones in the course or are poorly written, making it difficult to recover and forcing cleaning tasks and manual selection of publications.

Regarding the survey conducted among the participants of the MOOC, it should be noted that the population over 50 years old is that uses less and even acknowledged never do, compared to people aged 26-50. Regarding this, it is remarkable that those who are known as digital natives [319] and who are more familiar with these technologies are not the ones who use them most. Besides, there is not a high dispersion in the results regarding the perception of the utility of the hashtags, being this very positive and although with little difference, the group of digital natives is the one that less utility sees in the use of hashtags. Finally, also related to perception, those users who have not finished the course are those who, by far, find the use of hashtags less useful.

Moreover, it is clear how the Google+ network has made it possible to create a differentiated space for the course community through the Communities tool. Thanks to this separation is possible to quickly recover conversations even when users utilize their hashtags and not those defined in the course or directly without using them. On the contrary, in Twitter it seems essential to use in the tweets associated with the course some hashtag (non-formal) previously defined to detect the conversations in this way, otherwise, information is lost, and it becomes very difficult (although not impossible, according to Figure 27 and Figure 28) recovery of "informal" hashtags. So it is possible to say that Twitter makes difficult the recovery of informal conversations, unlike Google+.

Following on from the results can be remarkable an informal conversation within the community initiated by the students and especially associated to the contribution of content in the form of publication; the consequence of the non-formal activity initiated or even infused by the mechanics of the course. It is possible to verify how, for example, as a result of non-formal activities that are expressed as student publications, an entirely voluntary interaction of the rest of peers occurs through approval indicators such as "+ 1s", appearing more than 300 throughout the course associated with such non-formal activities or even associated comments.

On the results that compare the MOOC's ratings with performance in social networks, it is possible to affirm that in many cases, users who indicate their social network and who post messages in them have a more significant interest in completing the course. There are significant results that indicate a relationship between activities in both directions, although it is true that this relationship is more pronounced in the social network Twitter than in Google+.

As an outcome, and in conclusion, it is possible to affirm that it is possible to recover and classify non-formal and informal learning that students perform in environments such as social networks and that this knowledge can shed light on the complex learning processes that are given in multitudinous digital societies such as the one shown. In the case of this experiment, the results achieved in studying the conversations and interactions between users demonstrate the validity and usefulness of the software architecture developed.

As for other possibilities offered by this type of analysis, performing a more in-depth analysis and user level, it is presumed that it is possible to be able to classify types of users according to their activity in social networks and MOOC. The classification could be by the distribution over time, interest indicators, the possibility of increasing the segmentation for the adaptability of the MOOC platform through the data observed in Figure 27 and Figure 28, etc. This would allow researchers to find influential users, users who behave like spectators, users who do not have an interest in completing a course, but learning from the process, etc.

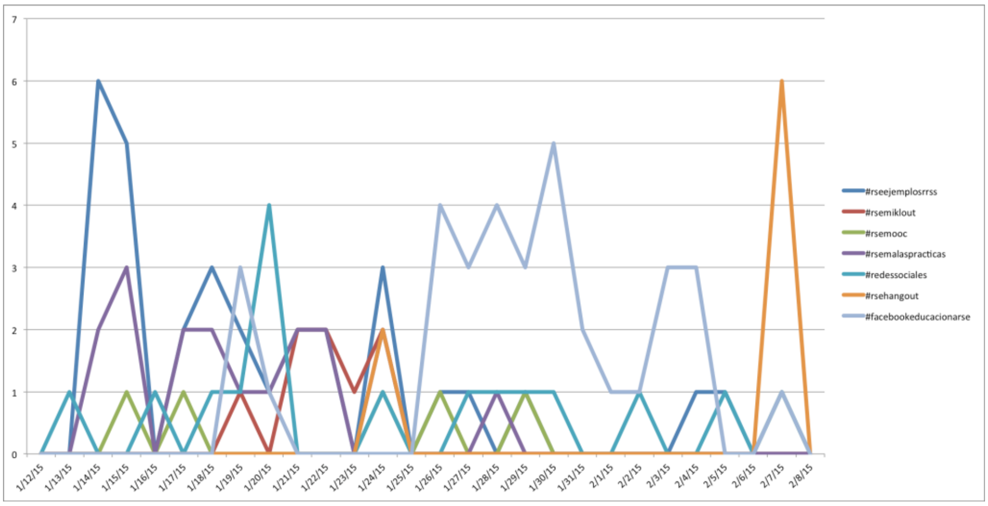


Figure 27. Evolution of the use of non-formal hashtags on Twitter throughout the course.

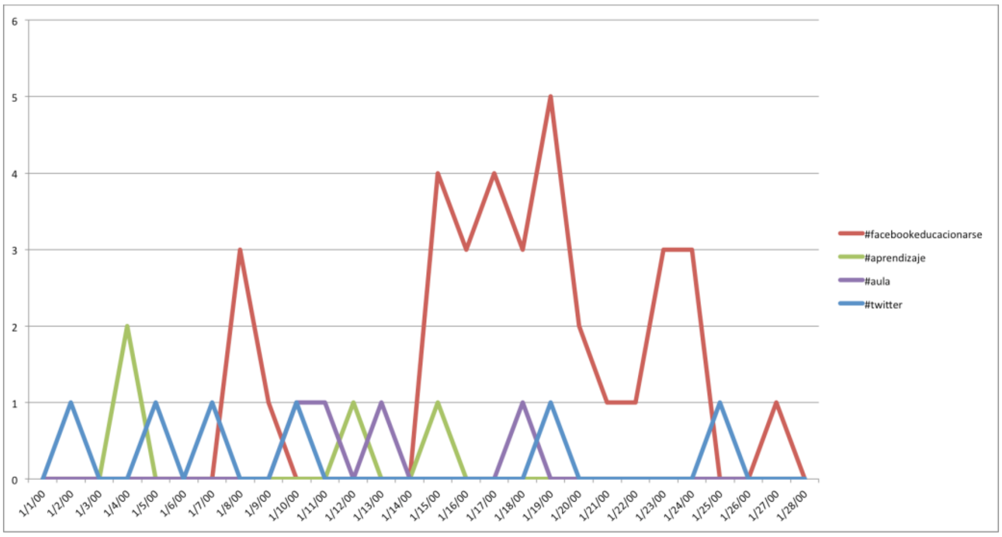


Figure 28. Evolution of the use of informal hashtags on Twitter throughout the course.

3.3 Analyzing and improving users' experience and performance in web forms

The collection of information using questionnaires and interviews is one of the most well-known and currently used methods to get users' opinions, both in the physical and digital environments.

It is common in many websites to have a form for entering information, either as a contact point, as part of the login for the system, as part of a payment process, etc. The forms are so integrated into the web user interaction, that their importance is relativized and it is assumed that the user completes it by the mere fact that they are faced with them regularly. However, this is not so.

Indeed, the web forms are pervasive: in the recent years have been triggered in the information collection methods certain trends and user behaviors towards such information entry tools. For example, it has been proven [60] the following regarding users' behavior towards forms:

- Users rely more on websites, even being more willing to perform complex actions (at all levels), such as purchases, payments, etc.
- Users protect more their information; they are less willing to disclose personal information.
- Users demand better products, are less tolerant to bad forms.

During the last years, much work has been carried out concerning the questionnaires, establishing that users have some reluctance to complete a form from even before to begin filling it [60]. This poses particular problems regarding the achievement of information collection objectives intrinsic to any form.

Regarding the types of users who complete forms, different profiles can be set [60]:

1. Readers: Those who read the form carefully.
2. Rushers: These users rush in and begin completing fields, reading only when they think it is necessary.
3. Refusers: These users won't have anything to do with the form.

According to the literature, and intimately related to the Social Exchange Theory [320], some authors [60] distinguish three layers in the forms: relationship, conversation, and appearance.

1. The relationship of a form is based on the relationship that who asks the questions has with whom responds.
2. The conversation of a form goes from the questions that are asked, to the instructions given or to the organization of the questions according to their topic.
3. The appearance of the form is the image it displays: placement of text, graphics, areas of data entry, color, etc.

Improving these factors, such as the relationship with the user, makes it easier for the user to participate and complete his task within the questionnaire.

The work presented in this section has been published in the papers [67, 321, 322] and the report [323] (published in the Spanish language). All of them are available in Appendixes 6.8, 6.9, 6.10, 6.11, and 6.12).

3.3.1 Context

Understanding what users do within a system is now a fundamental task in the digital world [324]. Most aspects of modern development workflows include users as a centric part of the design and development process of digital products (i.e., user-centered design [59, 325]). Not only knowing what users do (clicks, workflows, interactions, etc.) within a system is valuable for software developers and designers, but these stakeholders must also pay attention to other related aspects, like user experience, satisfaction, and trust [57, 63, 64, 326, 327]. Understanding what users do or feel when they use a system is extremely valuable to validate and improve a system. Analyzing users' interactions or their opinion about what they use makes it possible to ascertain the system's strengths or weaknesses regarding users' experience (mostly user interfaces and parts alike) to improve the system based on evidence.

Besides using the analysis of users' interactions and opinions to improve the worst-perceived parts of a system, developers can use these data to build custom or adaptive solutions for different kinds of users [61, 159, 328]. Using this idea, software engineers could develop versions of the system in which different version are shown to each kind of user. By knowing user profiles and identifying users' behavior and desires, the system could adapt its components to match users' expectations and likings better, and (probably) boost user performance and satisfaction [61, 327, 329].

For a better understanding, the context for this experiment is presented. The research has been conducted using a system that belongs to the Spanish Observatory for University Employability and Employment (OEEU in its Spanish acronym) [30, 330-333]. This observatory gathers data about employment and employability parameters among the Spanish graduates (after they leave the university) to analyze the information they provide and understand what the employment trends and most crucial employability factors are for this population. To accomplish this mission, the observatory has developed a complex information system [30, 333] that collects and analyzes data to present the insights to the researchers. The system is implemented using the Python language through the Django framework [231] and many other software libraries; it also keeps the information in a MariaDB relational database. To gather data from Spanish universities and students, the OEEU information system has two primary tools: one tool is devoted to obtaining students' raw data provided by the university; the other one is a system that generates custom web forms and questionnaires that are to be completed by the graduates after they leave the university. The problem of these web forms is their length, as they typically include between 30 and 90 questions. This second tool for gathering data (the questionnaires) is a centric part of this research.

During the months of June-July 2015, the Spanish Observatory for University Employability and Employment (OEEU) contacted several thousand Spanish university graduates (133588 individuals) through the universities (48, public and private institutions) where they got their degrees in the course 2009-2010 to invite them to fill out a questionnaire [330, 333].

This questionnaire had a shared part with 60 questions and 167 variables measured, in addition to 3 specific itineraries depending on the users' previous responses. The first itinerary added 3 questions and 9 measured variables more. The second one added 24 questions and 70 variables. Finally, the third itinerary added 32 more questions and 112 variables to the shared part of the questionnaire.

Therefore, the questionnaire varies between 63-92 questions and 176-279 variables depending on the itinerary that the user follows. It can be stated without doubt that the questionnaire is very extensive.

The number of users who started the questionnaire was 13006 (9.74% of the total population), of which 9617 completed it (7.20% of the total population, 73.94% of the total started questionnaires).

The descriptive data regarding the age of the participants in the questionnaire were the following (the count of users is 12109 because the birthdate data was not mandatory and not all users filled it out):

count	12109.000000
mean	32.525972
std	7.018282
min	25.000000
25%	28.000000
50%	30.000000
75%	34.000000
max	80.000000

As for gender, 56.05% (7290) of the users who answered the questionnaire were women and 43.94% (5716) men. About nationality, 98.54% (11672) of the users were Spanish, and 1.46% (173) were foreigners.

About the users who dropped out of the questionnaire, the quartiles of the dropout rate based on the questionnaire screen where they left off were:

count	3389.000000
25%	4.000000
50%	5.000000
75%	7.000000

That is, 25% of the users left on screen 4 or before, another 25% left between screens 4 and 5 of the questionnaire, another 25% between screens 5 and 7 and another 25% between screen 7 and the end (depending on the itinerary).

Moreover, in 2017, a process of gathering information similar to the one carried out during 2015 was conducted again. In this case, the information to collected was about graduates of masters studies that ended their studies during the 2013-2014 academic year. For this purpose, a questionnaire composed of between 32 and 60 questions and between 86 and 181 variables was proposed to gather data (the questionnaire has several itineraries again depending on the user's answers). Without too much analysis in detail, it can be considered that despite the differences, it is a lengthy questionnaire and shares some of the problems of the previous one regarding difficulties or challenges that can appear during its completion by the users.

Before sending out the questionnaires to the students, the Observatory gathers some data about students from the participant universities. In February 2017, 3 months before the public launch of the questionnaire for graduates of a master degree, there were collected data from 28744 people coming from 32 public and private Spanish universities. About these former students, the Observatory had the following data:

Descriptive data regarding the age of the population to which the questionnaire is addressed:

count	28744.000000
mean	35.854370
std	15.852381
min	5.000000
25%	28.000000
50%	31.000000
75%	38.000000
max	117.000000

Regarding the data about the age, the aging of the population concerning the one of the previous questionnaire was noticed. This is usually taking into account that the required age to begin a master degree is higher than that required to access to a degree (at least on a regular basis).

Regarding the gender of the population, 55.2% (16385) were women, and 44.8% (13317) were men. Concerning nationality, it was the aspect in which the population (graduates from a master degree) was more differentiated from the study performed with degree graduates. This time, the proportion of international students was more significant, with 88.11% (25318) of Spanish students compared to 11.88% (3414) students with foreign nationality.

In general terms, it was possible to assume that populations (putting each of them in context) were not very different. In this sense, can be highlighted the main difference was regarding nationality. This difference could lead to consider treating different aspects of the questionnaire to adjust to possible cultural differences. In this case, there was no cultural distinction when designing, presenting or performing the questionnaire. This could be considered a limitation of the study.

3.3.2 Goals

As presented in section 1.2, the objectives pursued in this scenario are:

1. Improve the user experience (and engagement) in a complex environment that involves a high amount of information and introduces a high level of friction for users when solving a task. This enhancement of UX seeks to improve the users' performance in the task proposed by reducing the friction.
2. Define an automated data-driven pipeline to analyze the users' interaction and improve the user experience.

To achieve these generalist goals in the case study presented, it is needed to define some specific goals that help to scaffold possible results and responses to the overall goals:

1. Design and validate different changes in the context of an extensive questionnaire regarding users' trust, user experience, usability and engagement with the final goal of improving the users' completion/success ratios.
    a. These possible improvements should be compared to the questionnaire developed previously by the Observatory for the same topics and context.
2. Present a new approach for enabling adaptability in web-based systems using A/B testing methods, user-tracking and machine-learning algorithms that could lead to improving user performance in completing a (large) web form, validating the obtained results through statistical tests.
    a. Produce all machine learning processes in a white-box way, using algorithms and techniques that allow researchers to understand what is happening in every moment.
    b. Moreover, to allow readers and other researchers to follow or reproduce the entire process, the author should provide all the

code used in the analysis process in Jupyter notebooks available publicly in Github.

3.3.3 Materials and methods

Materials and methods to design and validate the improvements in the questionnaire

The design process to improve the questionnaire was driven by a literature review (not as strict as an SLR). This literature review comprised about 650 books, papers, and technical reports. The process for selecting the literature to be reviewed was:

1. Making three different queries to the Web of Science and collect the results to iterate in reading the titles, abstracts and full content to select those papers really relevant for the topic of this research. The three queries performed were:
   a. *(("form*") OR ("questionnaire*") OR ("survey*")) AND "usability" AND "factor*" AND (("web") OR ("online"))*
   b. *online forms usability*
   c. *(("web" OR "online") AND ("questionnaire?" OR "form?") AND usability)*

   This process and its results are gathered in the following spreadsheet https://goo.gl/7n8xLg. In the spreadsheet, the 633 unique results retrieved from the Web of Science and their status regarding their usage in the research regarding each review stage are presented.
2. Extracting the main references from these papers and books retrieved from the Web of Science and read them. This process led to review another 15 papers, books, standards and technical reports. Most parts of them were used in some way to design the proposals that are explained below.

Once the literature review, the author designed the improvements and changes to the questionnaire. These improvements and changes are mainly supported or inspired by the literature as well as by ISO usability guidelines and HSS (U.S. Department of Health and Human Services) guidelines [326, 334-337].

The following subsections comment each change and measure, describing for each one its purpose, its goal, the identifier associated, etc.

After reviewing the literature to find what kind of changes should be applied to the questionnaire to be improved, increase the participation ratio and reduce the dropout, author prepared a list of the improvements proposal (available in [322, 323] and Appendixes 6.8 and 6.9). To validate the proposals designed, five experts were invited to evaluate the proposed measures using questionnaire. These experts were selected because all of them usually work with questionnaires from different perspectives (some of them work with questionnaires focusing on improving their usability, use them for research in several contexts, or design questionnaires as part of their day by day work). The assessment questionnaire completed by the experts was based on the proposal by Sánchez-Prieto *et al.* [104]. In it, the experts assessed the relevance of each proposed change, its clarity, and its importance, through a Likert scale (1-7 values). Also, the expert could comment on a qualitative way (typing comments in a textbox) any related issues to each question. Also, the questionnaire requires demographic data from the experts related to their gender, knowledge area, etc. [323] to characterize them. The cutoff point to accept each change or group of changes was defined in a Q1 (minimum 5.5 points out of 7 as maximum possible value).

Materials and methods to enable adaptability in web forms using A/B testing, user-tracking, and machine-learning algorithms

The questionnaires and custom web forms included in the OEEU information system gather data from students in two ways: information provided explicitly by the students (the information provided directly) and *paradata* [338]. The paradata from these questionnaires are the auxiliary data that describe the filling process, such as response times, clicks, scrolls, and information about the device used when using the system. All the data used in this part of the research are taken from these two available sources: the raw input tool used by universities and the web forms tool (providing user inputs and their paradata).

Regarding the data, it is worth noting that to characterize the main factors that affect users in completing the questionnaires, the data chosen have been only those available before the users began the questionnaire. This is because the research is focused on investigating which factors predetermine participants' success or failure in completing the form, considering all the factors related only to personal context and device and software used to access the web forms. The data about the personal context of the user are provided by the OEEU's system and include information submitted by the university where the user (graduated) studied. All the information that could be used to create the models that predict whether the user will complete the questionnaire (before starting it) is presented in Table 19. Table 19 also explains the data variables used and whether they were valuable for the models. This research has been carried out with a total of 7349 users (all who have some experience with the web forms during the experiment). Of them, the data from 5768 users were considered initially. Finally, data from 3456 users (those resultant after cleaning the data) were used to train and try the machine-learning algorithms and perform the statistical tests; 1165 users were the cohort introduced in a phase of reinforcement for the questionnaires. This reinforcement phase was used to validate the rules generated to adapt the web form to users. This number (1165) included users who did not complete the web form in the first stage as well as users that joined the experiment during the reinforcement and validation phase. Other users (416) only viewed the web forms without starting them. For that reason, were not considered in the experimental report.

As found in the bibliography (and as a result of the previous part of the case study) the concept of A/B testing —also known as bucket testing, controlled experiment, etc. — applied to websites and the Internet could be explained as follows: "show different variations of your website to different people and measure which variation is the most effective at turning them into customers (or people that complete successfully a task in the website, like in this experiment). If each visitor to your website is randomly shown one of these variations and you do this over the same period, then you have created a controlled experiment known as an A/B test" [339-341]. In this case (and again as result of the previous part of the research), the author has prepared three different variations, called verticals A, B, and C. In each variation, the author, introduced several changes related to enhancing the users' trustiness, engagement, make the user interface more conversational, etc. All these changes, introduced in the different variations of the web forms (the verticals) used in this research, were proposed [322]. These verticals were used as the website variations in which users (students responding to the questionnaires) are meant to test which version is the best regarding the users' performance in the initial stage. To do so, before the experiment, 5768 users were redirected randomly to the different vertical. In the last part of the experiment, the verticals were used to check whether the rules and users' analysis performed during the machine-learning analytics process improve the users' performance in completing the web forms. In this validating phase (which also is called reinforcement in this section), 1165 users were redirected to the verticals using the rules generated analyzing the interaction data from the users that acceded randomly to the verticals.

Table 19. Initial variables gathered from the OEEU information system to build the predictive models of questionnaires' completion.

Name of the variable in the code	Explanation	Type of information that it provides	Was this variable used finally to build the predictive models?
estudiante_id	ID number of student	Personal information	Yes
annoNacimiento	Year of birth	Personal information	No
sexo_id	Gender (male / female)	Personal information	Yes
esEspannol	Is the student Spanish?	Personal information	No
universidad_id	ID of the university where the graduate studied	Personal information	Yes
estudiosPadre_id	Maximum educational level achieved by the graduate's father	Personal information	No
estudiosMadre_id	Maximum educational level achieved by the graduate's mother	Personal information	No
situacionLaboralPadre_id	Current employment status of the graduate's father	Personal information	No
situacionLaboralMadre_id	Current employment status of the graduate's mother	Personal information	No
oficioProfesionPadre_id	Occupation of the graduate's father	Personal information	No
oficioProfesionMadre_id	Occupation of the graduate's mother	Personal information	No
residenciaFamiliar_id	Place of residence of the graduate's family	Personal information	No
residencia_id	Place of residence of the graduate during studies	Personal information	No
idMaster_id	ID number of the master study	Personal information	Yes
especializacionMaster_id	Specialization of the graduate's master	Personal information	No
masterHabilitante	Is an enabling master?	Personal information	No
titularidadMaster_id	Public or not master	Personal information	No

Name of the variable in the code	Explanation	Type of information that it provides	Was this variable used finally to build the predictive models?
modalidadMaster_id	Modality of the master (online, physical, etc.)	Personal information	No
cursoInicioMaster	Season of the beginning of the master	Personal information	No
cursoFinalizacionMaster	Season of the completion of the master	Personal information	Yes
notaMedia_id	Average grade of the student	Personal information	No
realizacionPracticasMaster	Did the student professionally practice during the master?	Personal information	No
tiempoDuracionPracticasMaster	Time spent by the student in professional practices during the master	Personal information	No
realizacionErasmusMaster	Did the student do an Erasmus stay?	Personal information	No
tiempoDuracionErasmus_id	Time spent by the student in an Erasmus stay	Personal information	No
paisErasmusMaster_id	Country where the student did an Erasmus stay	Personal information	No
viaAccesoMaster_id	Way of accessing the master	Personal information	No
verticalAsignado	Vertical assigned in the A/B testing for the student	Experiment configuration	Yes
cuestionarioFinalizado	Did the student finalize the questionnaire?	Experiment configuration	Yes
numUniversidades	Number of universities involved in the master	Personal information	Yes
numUniversidadesEspannolas	Number of Spanish universities involved in the master	Personal information	Yes
ramaConocimiento_id	Knowledge branch of the master (healthcare, social sciences, engineering, etc.)	Personal information	Yes
realDecreto	Official statement approving of the master studies program	Personal information	Yes
browser_language	Language of the browser used	Device information	Yes
browser_name	Name of the browser used	Device information	Yes
browser_version	Version of the browser used	Device information	Yes
device_pixel_ratio	Device pixel ratio of the browser	Device information	Yes
device_screen_height	Device screen height	Device information	Yes
device_screen_width	Device screen width	Device information	Yes
landscape	Is the device in landscape mode?	Device information	No
os	Operative system of the device	Device information	Yes
os_version	Version of the operative system used	Device information	Yes
portrait	Is the device in portrait mode?	Device information	No
push_notification	Did accept the graduate push notifications for the web form?	Device information	No
push_notification_id	ID number for the push notification subscription	Device information	No
tablet_or_mobile	Is the device tablet or mobile?	Device information	Yes
userAgent	User agent of the device used	Device information	Yes
viewport_height	Height of the window browser	Device information	Yes
viewport_width	Width of the window browser	Device information	Yes

The variables excluded from building the predictive models are those that have more than 10% of their observations with the null value.

The programming language used to conduct all the analyses and calculations in this phase was Python. The Python software tools and libraries used to code and execute the different algorithms and statistics were:

- Pandas software library [342-344], to manage data structures and support analysis tasks.
- Scikit-learn [345] library, to accomplish the machine learning workflow [346].
- Jupyter notebooks [347-349], to develop the Python code used in this research.

All the code developed to analyze user interactions and create machine-learning models etc. is available at https://goo.gl/Zy5WZp [350].

In general, the performed analysis (based on statistics and machine learning) follows universal principles in data science regarding data structuration, tidy data approaches, etc. [31, 342, 344]. The machine-learning process, as commented in the goals section, were implemented in a white-box way; thus, the author has selected algorithms and methods to make the workflow explainable. This is extremely important, from the author's point of view, in a research project like this, as it allows humans to provide feedback to the algorithmic process.

Materials and methods for testing how different versions of layout and complexity of web forms affect users

In the case of this part of the experiments, the research was conducted analyzing 123 users involved in the experiment. These 123 users were selected because they were swapped between different versions of the form in the reinforcement phase (after applying the changes using A/B testing).

Analyzing this sample, the author studied the users' performance related to each version of the web form and swapping users between both versions to test what is the effect of this change in their performance.

To do so, it is proposed a quasi-experimental research design with a control group. Following this design, the users were divided into two groups: the experimental group (89), composed by the users that were redirected from one questionnaire design to a different one, and the control group (34) composed by those users that remained in the same questionnaire design.

After the application of the different treatment to each one of the groups was compared the differences in the finalization rate using three-dimensional-contingency tables and chi-squared to analyze the impact of increasing or decreasing the complexity of the questionnaire. In consequence, the following hypotheses were posed:

- H1. The redirection to a different version of a questionnaire has an impact on the finalization rate.
- H2. The redirection of users from a text plain questionnaire to one with more complex elements has an impact on the finalization rate.
- H3. The redirection of users from a questionnaire with complex elements to a plain text one has an impact on the finalization rate.

3.3.4 Results

In this subsection, the results of each phase of the research are commented on separately.

Results of designing and validating the improvements in the questionnaire

The changes proposed for the questionnaire were:

1. TR1. Modify the text and appearance from the invitation letter to the questionnaire.
2. TR2. Adequacy of the image to the other digital products of the Observatory.
3. TR3. The inclusion of the Observatory's logo and university's logo.
4. TR4. Changes in the introduction text to the questionnaire.
5. US/UX1. The inclusion of a progress bar in the questionnaire.
6. US/UX2. Present a visual focus animation on concrete actions.
7. US/UX3. Deactivation of control elements when an action is initiated.
8. US/UX4. In related elements, instead of having smaller and more specific groupings, use some larger grouping, following the Gestalt principles on grouping.
9. EN1. In the questions related to the community in which they live, change the drop-down selector for a map with the autonomous communities of Spain.

10. EN2. Inclusion of textual feedback related to user responses including information that may be relevant.
11. EN3. Inclusion of web push notifications that allow Observatory to send messages to users to encourage them if they leave the questionnaire before finishing.

Table 20. Relationship between each change/improvement proposed, the HCI application areas and the layers of Social Exchange Theory [320]. Source: [322].

Layer of the Social Exchange Theory	Relationship	Conversation		Appearance
Improvement area regarding HCI	Trust	Engagement	Usability / User Experience (UX)	Design
TR1	X			X
TR2	X		X	X
TR3	X			
US/UX 1			X	X
US/UX 2			X	X
US/UX 3	X		X	
US/UX 4			X	X
TR4	X	X		
EN1		X		X
EN2		X		X
EN3		X		

Table 21. Descriptive results from the experts' evaluation for each proposal regarding the pertinence, relevance, and clarity.

	Pertinence			Relevance			Clarity		
	AVG	STD	N	AVG	STD	N	AVG	STD	N
TR1	6.17	0.98	5	5.17	1.17	5	6.17	0.98	5
TR2	6.17	1.33	5	6.00	1.55	5	5.67	1.37	5
TR3	6.67	0.52	5	5.83	1.47	5	6.67	0.52	5
TR4	6.33	1.03	5	6.00	0.89	5	6.50	0.84	5
US/ UX 1	6.50	1.22	5	6.83	0.41	5	6.50	1.22	5
US/ UX 2	7.00	0.00	5	7.00	0.00	5	6.67	0.52	5
US/ UX 3	5.67	2.42	5	6.00	2.45	5	5.33	2.42	5
US/ UX 4	7.00	0.00	5	6.67	0.52	5	6.67	0.52	5
EN1	6.83	0.41	5	6.33	1.03	5	5.67	1.03	5
EN2	4.83	2.14	5	5.50	1.22	5	5.17	2.32	5
EN3	7.00	0.00	5	7.00	0.00	5	7.00	0.00	5

Table 22. Descriptive results from the experts' evaluation for each group of proposals and global assessment regarding the pertinence, relevance, and clarity.

	Pertinence			Relevance			Clarity		
	AVG	STD	N	AVG	STD	N	AVG	STD	N
TR	6.24	1.00	20	5.76	1.26	20	6.24	1.04	20
US/UX	6.48	1.47	20	6.62	1.32	20	6.24	1.51	20
EN	6.19	1.64	15	6.25	1.13	15	5.88	1.67	15
Global	6.27	1.30	55	6.22	1.30	55	6.11	1.42	55

To get a full description of each change and how they can be implemented in the web form, please refer to [322] or Appendix 6.8. The relation of each improvement proposal and the HCI area and the layer of the Social Exchange Theory are presented in Table 20.

The results of the experts' assessment of each change proposed, and according to the evaluation method described in materials and methods section, are presented in Table 21 and Table 22.

In general, the average mark of the assessment in each question and grouping topic could be considered as good: most of the results are in the Q1 (score 5.5).

This Q1 score is not achieved in the proposed change EN2 (*inclusion of textual feedback related to user responses including information that may be relevant*) pertinence and clarity, TR1 (*modify the text and appearance from the invitation letter to the questionnaire*) regarding its relevance and the US/UX3 proposed change (*deactivation of control elements when an action is initiated*).

In these cases, all the evaluations exceeded the Q2 score (4.0 value), so still they could be considered as well perceived changes but, in any case, they were reviewed again by the researchers, to improve the final version of the online form.

Results of enabling adaptability in web forms using A/B testing, user-tracking, and machine-learning algorithms

The workflow resulting from this research (available at https://goo.gl/Zy5WZp [350]) was as follows:

1. Retrieve datasets about users from OEEU's information system.
2. Filter the desired fields from the datasets and merge datasets in a single data frame (a data structure like a table).

3. Data cleaning: remove noise data, remove columns (variables) with too many null (*NaN*) values, and remove all users who have only partial information and not all presented in Table 19.
4. Normalize data with the One-hot encoding algorithm for categorical values in columns [346]. To apply the One-hot encoding, researchers used the *get_dummies()* function from Pandas library, as presented in [350].
5. Considering the data gathered and the kind of variable (labeled) to predict, the algorithm to use must be related to supervised learning. This is because this kind of algorithm makes predictions based on a set of examples (that consist of a labeled training data set and the desired output variable). Moreover, regarding the dichotomous (categorical) character of the variable to predict, the supervised learning algorithm to apply must be based on classification (binary classification, as it existed a label of finalization equal to *true* or *false*). According to the author's previous experience, the possibility of explaining results and the accuracy desired for the classification, a Random Forest classifier algorithm [351] was selected. In this step, the Random Forest algorithm was repeatedly executed, using a custom method based [350] on *GridSearch* functions from Scikit-learn, to determine the best setup for the dataset given (obtaining the most valuable parameters for the execution).
6. With the best configuration found, train the random forest algorithm (with 33.33% of the dataset) and obtain the predictive model.
7. Using the predictive model, obtain the most important features for the predictive model. To obtain these features, the author applied *feature_importances_* method from the Random Forest classifier implemented in Scikit-learn library [350].
8. Using the most important features (those that have importance higher than a custom threshold value of 0.05—the importance score varies between 0 and 1, where 0 is the worst score and 1 the best one), generate clusters applying hierarchical clustering [352]. The reason to use hierarchical clustering is that the algorithm does not require

deciding upon the number of clusters to obtain (so, it does not also require to fix Euclidean distances and other parameters previously); it obtains all possible clusters showing the Euclidean distance between them. These clusters represent the groups of users who have participated in the questionnaire according to the most critical factors found in the classification.

9. With these clusters, the researchers investigate which clusters exhibit low performance.
10. Using this knowledge about groups of users with low performance and the heuristics observed, software engineers responsible for the OEEU's information system and its web forms could propose changes and fixes (rules, redirections, etc.) in the platform that might help users to improve their performance in the future.
11. Once the data-gathering process is finished, the researchers performed a statistical analysis of the finalization rate of the individuals to determine whether the application of the rules had any impact in the improvement of the finalization of the questionnaires. With this purpose, and considering the characteristics of the variables, the author applied the Chi-squared test given that it is the most convenient alternative for the analysis of the relationship between two nominal variables.

All these steps are summarized in the Figure 29.

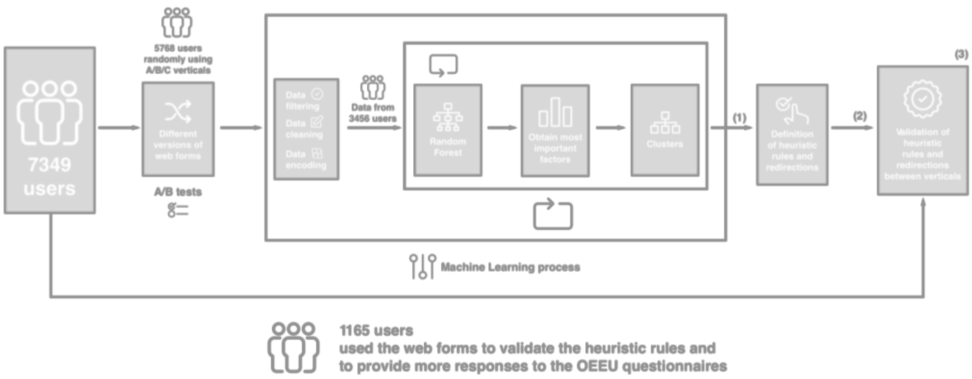


Figure 29. Overview of the process followed to enable adaptability in web forms.

As the first result of this experiment, the workflow designed was implemented in Python using the libraries and tools previously commented. Also, it was publicly released in Github.

Concerning the validity and utility of this workflow, there are two main results to highlight:

First of all, using the workflow established was possible to generate predictive models on what are the main factors that have an impact on the finalization rate of the users. The predictive models that performed better and were used for the experiment were three: one per vertical (A/B/C versions of the questionnaire). Each model described the most influential factors for the users to complete the questionnaire. In the following tables (Table 23,

Table 24, Table 25, Table 26, Table 27, Table 28) are described the results of the predictive models per each form vertical.

Table 23. Results of the predictive model for the vertical A.

	Precision	Recall	F1-score	Support
False	0.92	0.35	0.51	69
True	0.85	0.99	0.92	263
Avg / total	0.87	0.86	0.83	332

Table 24. Results of the predictive model for the vertical B.

	Precision	Recall	F1-score	Support
False	0.92	0.37	0.52	161
True	0.74	0.98	0.85	301
Avg / total	0.81	0.77	0.73	462

Table 25. Results of the predictive model for the vertical C.

	Precision	Recall	F1-score	Support
False	0.89	0.37	0.52	132
True	0.74	0.97	0.84	238
Avg / total	0.79	0.76	0.73	370

Table 26. Crosstab of the predictive model results for vertical A.

	False (predictions)	True (predictions)
False (actual)	24	45
True (actual)	2	261

Table 27. Crosstab of the predictive model results for vertical B.

	False (predictions)	True (predictions)
False (actual)	59	102
True (actual)	5	296

Table 28. Crosstab of the predictive model results for vertical C.

	False (predictions)	True (predictions)
False (actual)	49	83
True (actual)	6	232

In this case, and according to the goal of producing the ML research using white-box models, was possible to extract the relevance of each factor in each predictive model generated. The results were the following:

Most influential factors for the predictive model for vertical A:

1. *viewport_width:* 0.267931
2. *tablet_or_mobile:* 0.139438
3. *os_iOS:* 0.132425
4. *device_screen_height*: 0.118814
5. *device_screen_width:* 0.067581
6. *device_pixel_ratio:* 0.066577
7. *os_Android:* 0.054088

Most influential factors for the predictive model for vertical B:

1. *viewport_height*: 0.294176
2. *viewport_width:* 0.167701
3. *device_screen_height*: 0.102463
4. *device_pixel_ratio:* 0.085122
5. *os_Android*: 0.076196

Most influential factors for the predictive model for vertical C:

1. *device_screen_width*: 0.193903
2. *viewport_height:* 0.143456
3. *device_screen_height:* 0.108721
4. *tablet_or_mobile:* 0.100000
5. *viewport_width:* 0.093584
6. *device_pixel_ratio:* 0.088479
7. *os_Windows:* 0.055153

Using the most relevant factors, the author generated clusters of users (based on similarities between them) and a set of heuristic rules that summarized what the differences between the different verticals in the case of how the different factors affect the users' performance on finalizing the questionnaire in each case are. Using these heuristic rules, in the reinforcement phase of the questionnaire, users were redirected (knowing about them the information shown in Table 19) to the vertical that could be the best for them (concerning performance). Also, after analyzing the results in the first phase of the questionnaire, was decided to dismiss the vertical C of the questionnaire, since it did not present relevant improvements over the other two verticals.

After completing the experiment, the results achieved were promising. Using the heuristic rules and redirecting users to the version that fits better for them worked and impacted positively on the users' completion rate (69.34% previously, 71.59% after the reinforcement with the redirection). Also, comparing the users' performance in completing the questionnaire between the different redirections done, all of the redirection shown a positive impact. The complete results are available, including analyses at heuristic-rule-level, in [67] and in Appendix 6.10.

Overall results of UX and users' performance after improving the questionnaire and use A/B testing

The results achieved after using the three-dimensional-contingency tables were:

- Reject hypothesis 1 (H1). So the redirection of users between different versions of the questionnaire had no impact on users' finalization rate.

- Accept hypothesis 2 (H2). The statistical tests support that the redirection of users from a simpler questionnaire to a more complicated one has an impact on the finalization rate.
- Reject hypothesis 3 (H3). The redirection of users from a more elaborate questionnaire to a simpler one has no impact on the finalization rate.

This part of analyzing statistically each issue that affected the finalization rate of users is not finalized. Author plan to continue this research to delve deeply into the results raised during all the experiment.

More information about these statistical tests and the work done is available in [321] and in Appendix 6.11.

3.3.5 Discussion

It is worth to comment, first of all, that the A/B testing approach used for this research is not a pure application of such methodology. While A/B tests are commonly based on singular changes between the different experimentation groups (or verticals), in the presented approach the author grouped different changes into the same verticals. In this case, this variation of A/B tests does not influence this experimentation, as the researcher(s) attempts to maximize user performance in the questionnaire finalization without a particular focus on small changes, but using essential differences between the different verticals. Despite that, it is worth noting that this kind of application of A/B tests for the experiment has been previously validated by experts [29].

Is it advisable to apply this kind of machine-learning method to this kind of problem? In this case, the researchers (the author of this thesis and collaborators) were inspired by other authors who have applied these types of processes to a wide range of problems. As an example, this kind of machine-learning algorithmic approach has been used in other fields, such as education [353], with promising results. Beyond the benefits that machine-learning approaches bring to many problems, by also including white-box procedures, explainable and reproducible results that could be improved or discussed by the scientific community are ensured. All these considerations and precedents encouraged the author to employ this kind of approach to address the problem of improving users' performance within a complex system like that presented. According to the results, the question can be answered positively, as the findings have been valuable and proved the validity of the approach.

Regarding the generated predictive models, the cut-off value for their relevant factors to later include in the clusters; the author stated 0.05 as the minimum value to consider since this is the most common value in classical literature to ensure reliable results. Also, in this case, the author uses this cut-off value to generate the clusters using only the most important factors (those that have a specific weight of more than 0.05 in the predictive model), thus excluding less important ones that could introduce noise when building the groups.

Concerning the most important factors that characterize the predictive models and explain the users' profile and preferences while completing the questionnaire, it should be remarked that technical aspects were more important than personal ones. At the beginning of the research and the predictive models' generation, researchers included personal aspects, such as gender, age, and issues related to education, as part of the dataset. According to the results, such aspects do not have special relevance while modeling the users' behavior in completing the web form. Instead, the present findings indicated that the most important factors for the users were the size of the device screen and the browser window. Moreover, other aspects, like the screen resolution, specific browser, or operative system, were necessary, but with a lesser effect. Nevertheless, these are the most important factors for the population of this study and cannot be considered general and valid for other populations. To apply the approach presented in this research in other experiments, the predictive models should be generated again.

Regarding the generation of rules based on heuristics, and as a future study, the researchers involved in the experiment would like to automate this process. This will help to reproduce the same process with the same experimental conditions and remove any bias introduced by researchers or administrators. This is explained in depth in the following subsection.

Related to the reinforcement phase and other conditions of the experiment, with the aim of enhancing users' participation in the questionnaires, the OEEU offered participation in a raffle (the prize would be seven smartwatches) to all graduates completing the web form as a reward. This incentive was also used to promote the reinforcement process where the redirection rules were applied.

Regarding the effectiveness of the use of rules based on cluster analysis during the reinforcement period, cluster analysis was found to be a handy tool to guide the redirection of users to the version of the questionnaire best suited to the features of the technology with which they completed it.

Another interesting future line of research would be an analysis of the threshold cut-off to perform the factor selection, given that a higher minimum value may simplify the number of rules and make more efficient the redirection process. As a first step, the author intends to analyze rule four to gain a better understanding of the predictive importance of the elements behind its formulation.

Finally, the author believed that the approach and procedures presented in this research are transferable to other application fields. The process presented in this research follows some traditional approaches and methods within the machine-learning research field, and the prediction challenge is present in many other problems beyond web form completion. The proposed methodology may also help to transfer this experience to other problems with the added value of providing a white-box approach for the algorithms used. For that reason, the author wanted to attempt to apply such methodology to predict the employability of Spanish graduates.

In this case, after the experiment, the author employed the same process resulting from this research. The resulting paper of the application shown positive results, validating the white-box procedure. The new experiment using the same workflow is available in [354] and in Appendix 6.12.

3.4 Aiding programmers in Quantum programming

Quantum computing (QC) programming is not currently an easy task. New languages like Open QASM [355] and SDKs like QISKit [356, 357] open new horizons for the research and development within the new paradigm of quantum computing [358, 359]. Despite that, these new languages and SDKs present a steep learning curve for regular developers and newcomers in the field of quantum computing. Aiding programmers by enabling recommenders, assistants or other intelligent agents can make this learning curve smoother and help to popularize quantum computing, at least regarding code development. This approach, i.e., using new ways of teaching how to code, is not new; Seymour Papert in 1984 [360] referred to this concept (using *expert human* teachers instead of intelligent helpers) as: "[...] fluency in programming provides an opportunity for teachers to teach in new ways and for students to learn in new ways" [360].

In this case, the environment used for the research is more controlled than for the previous experiments/scenarios. The usage of a language such as QASM or other programming languages implies that the user should follow standard rules so that the programming results can be seen as a chained probability function. In this chained probability function, the output is the result of the weighted probabilities of each occurrence of a reserved character or word in a sentence or program line to compose a full program and produce a kind of result. More specifically, programming could be viewed as a "(...) model of a system which produces such a sequence of symbols governed by a set of probabilities". When considering this definition, it is possible to assume that programming is a stochastic process. A lot has been written about stochastic processes in several areas of science, such as Physics [361]. Related to the definition, the transmission of information to be processed by the computer is key, and the concepts explained by Shannon in his famous work "A Mathematical Theory of Communication" [362] are adopted in this experiment.

Based on the view of programming as a stochastic process, the experiment goal is to provide an intelligent system capable of helping programmers that use a quantum programming language —Open QASM in this case— to write code in a more accessible mode.

The work presented in this section has been published in [363] (paper available in Appendix 6.13) and can be considered a work in progress. As discussed below, although testing with real users and formal evaluation is still needed, the initial results are promising and encourage the author to continue the research, improving the technical solution achieved in order to deploy it in a real system.

3.4.1 Context

Nowadays there are many ways to build intelligent systems that can learn from humans and build a knowledge corpus on their own. These knowledge corpora could be used to provide a different kind of help to humans in tasks like aiding in decision-making processes, recommending multimedia resources, building conversational agents, etc.

Related to the aforementioned intelligent systems, in the literature and new media appear buzzwords like Artificial Intelligence (AI) [364], Machine Learning (ML) [345, 346], Deep Learning (DL) [365, 366], etc. Many modern applications include these kinds of concepts and keywords to look trendy; others involve them to deal with problems that are difficult to solve in other traditional ways. Apart of the trendiness of the terms, it is a fact that these research fields are increasingly present: many enterprises are spending a lot of effort and money to be AI-driven; including AI in applications, decision systems, etc. In this sense and related to the concept of aid provided by AI or intelligent systems, it is presented the core-concept of User Experience (UX). The UX is commonly defined as “a person’s perceptions and responses that result from the use or anticipated use of a product, system or service” [367]. Related to the ISO definition, UX includes all the users’ emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviors and accomplishments that occur during, before and after the usage. According to this ISO, three factors influence user experience: system, user and the context of use. Considering these three factors, and related to the new wave related to AI, it could be interfered on the system or the context of use [334] by applying AI to existing systems. Thinking in the AI application, the UX can be improved since the AI could learn from users' and previous usage to change or adapt the system or specific features to the current user's needs, desires and behaviors.

Many times, the UX is merely considered in the context of visual applications (visual UIs) or related to common products designed for regular users. In the case of this work-in-progress research, it is tried to improve the user experience in the context of programming under the quantum computing paradigm. Learning how people to code using the products developed by IBM Research [368], the assumption is that is possible to distill knowledge to later use it in guiding and helping other quantum-programmers in common tasks related to the code. Based on the experience gained in this field, currently there are many common issues on coding quantum programs (definition of qubits to use, measurement operations, etc.) based on some standard rules and patterns, that could represent general (and simple) cases where the users could initially be helped. This help could positively affect the users by fulfilling their desires and expectations to achieve success with their code and experiments or on creating more positive experiences through being helped by an intelligent system that could provide real-time feedback on their code [57, 369].

Gathering all these ideas and core concepts, in this scenario is described a work-in-progress (WIP) project developed by the author within the AI Challenges & Quantum Experience Team. This project implements an intelligent system based on a deep learning approach that learns how people to code using the Open QASM language to offer help by recommending different code sequences, logical steps or even small pieces of code.

3.4.2 Goals

As presented in section 1.2, the objectives pursued in this scenario are:

1. Study and propose a way to aid users in solving an incredibly complex task, even when they do not have enough knowledge to solve it.

2. Secondary goal: based on the proposal raised from the primary objective, create software that helps to aid users and could be integrated into different environments related to the problem.

These goals can be specialized in the case of this case study to:

1. Offer help to quantum computing programmers by recommending different code sequences, logical steps or even small pieces of code. The recommendations should be based on the code typed previously by the programmer.
2. Implement an intelligent system based on a deep learning approach that learns how people to code using the Open QASM language to recommend code to programmers as described in the previous goal.

3.4.3 Materials and methods

What is source code? In fact, and with no intention of providing a broad definition, the code is a human-readable set of words or alphanumerical characters (instructions) previously defined that could be accompanied by other characters like punctuation, etc., and follows a logical structure and some kind of grammar [370, 371]. Following this consideration (and the *programming languages* idea), the foundations of coding are not so far from those that define the human languages.

In the Natural Language Processing (NLP) area, many researchers work using artificial intelligence to analyze human language and design conversational systems, to summarize texts automatically, to learn and replicate communicative styles, etc. In this sense, this research uses concepts from NLP to teach neural networks how to code using quantum computing languages and libraries (mainly using Open QASM [355]) like the human programmers. That is, through feeding Recurrent Neural Networks (RNN), the code entered by programmers when they use the IBM Q Experience [372] to enable them to learn how the code is composed in the context of quantum programming.

The approach followed is under the NLP umbrella: it is the sequence-to-sequence (*seq2seq*) neural network model [373-375]. This model consists of two RNN's: "one RNN encodes a sequence of symbols into a fixed-length vector representation, and the other decodes the representation into another sequence of symbols. The encoder and decoder of the proposed model are jointly trained to maximize the conditional probability of a target sequence given a source sequence" [374]. That is, using this method is possible to train a system that produces sequences of symbols using an input sequence of symbols (Figure 30).

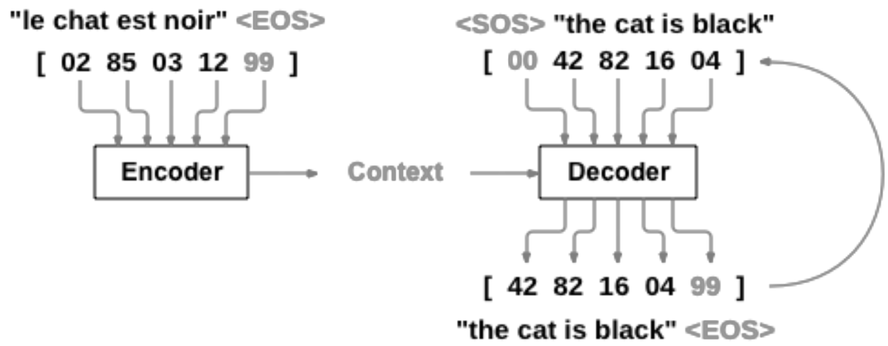


Figure 30. Overview of a sequence to sequence network using different networks as encoders and decoders. The image is taken from [376].

This approach has been used recently with success in the tasks of performing translations between different languages [377, 378] (Figure 30), since its ability to learn semantically and syntactically meaningful representation of linguistic phrases [374].

In this case, the *seq2seq* neural network is not employed to translate languages, but it is used similarly to translate some sequences to other ones. In this approach, the input sequence is the one typed by the user in the tools for developing quantum code, and the target sequence is the next logical sentence(s) proposed by the neural networks.

3.4.4 Results

Based on previous experiences by the author [67] and other researchers [369, 379], the main idea of this research is that intelligent systems could enhance the user experience. Furthermore, they could enhance the experience of developers while they code new programs in challenging environments like programming in Quantum Computing-specific languages. As previously stated, the goal is to provide real-time feedback to IBM Q users by proposing code to them in the different quantum programming environments developed by IBM Research and IBM Q Experience team [372]. To provide the recommendations to the users, is being developed a neural network based on a *seq2seq* approach that learns the sequences from the code composed by the users to develop quantum programs in both ways: learning what kind of sentences are used (and relevant for quantum computing) and also the logical sequences they follow to build a quantum program.

To pursue this, there are some critical factors to keep in mind: how to develop the *seq2seq* network and its different utilities and how to train it using the data available in IBM's platforms.

First of all, regarding the technological issues, the deep learning framework selected to develop the solution is PyTorch [380], supported by the *PyTorch-seq2seq* [381] library from IBM, which is used to build the *seq2seq* network. PyTorch has been selected due to its features related to the creation of dynamic neural networks and its performance in building deep learning models using GPUs. Also, the *seq2seq* library for PyTorch developed by IBM is used to work from a previously tested approach regarding the *seq2seq* models and to avoid starting from scratch. This library encloses different functionalities related to the RNN that encodes the input sequences and the decoder RNN that produces the output (target) sequences. Apart of the functionalities, it includes, it has been developed a variation (available in https://goo.gl/KFmaKW) on the prediction method used to produce the target sequences. This variation on the prediction method not only produces a unique prediction based on the most probable output sequence (as in the original implementation) but also includes a *beam search* strategy [373, 382] to produce several possible outputs (target sequences) with variations given an input sequence. In fact, with this change, it has been replaced with the default RNN decoder to use a TopKDecoder RNN.

Secondly, the author developed datasets to train the models in the task of predicting the proper code sequence that would be the result of adding the next logical sentence to the given code sequence. The datasets have been built using the following rule *n -> n+1:* given a code sequence 'n', the predicted sequence would be 'n+1', where 'n' is the set of instructions that compose the input sequence, and the '+1' is the following instruction that could be used after the last instruction in the set 'n'.

The dataset development consisted of two main phases: (1) designing the code sequences to represent inputs and outputs (given and target sequences) and (2) building simplistic representations of the code sequences to facilitate the training. On the designing of the code sequences, the original dataset of quantum programs sent to the IBM Q backends was augmented by dividing the quantum program into all the possible parts that followed the rule *n -> n+1 (input -> output sequence)*. To reduce the complexity of the training phase, each different instruction was replaced using a unique key, thereby mapping the instructions used in the code developed by the users to get more straightforward representations. This unique key (composed of 1-3 alphanumerical characters typically), is used in the datasets to train the encoder and decoder RNN. To build the mapping, all the defined Open QASM statements and their possible arguments were parsed. The list of these Open QASM statements is available in Figure 31.

<table>
<tr><th>Statement</th><th>Description</th><th>Example</th></tr>
<tr><td>OPENQASM 2.0;</td><td>Denotes a file in Open QASM format</td><td>OPENQASM 2.0;</td></tr>
<tr><td>qreg name[size];</td><td>Declare a named register of qubits</td><td>qreg q[5];</td></tr>
<tr><td>creg name[size];</td><td>Declare a named register of bits</td><td>creg c[5];</td></tr>
<tr><td>include "filename";</td><td>Open and parse another source file</td><td>include "qelib1.inc";</td></tr>
<tr><td>gate name(params) qargs</td><td>Declare a unitary gate</td><td>(see text)</td></tr>
<tr><td>opaque name(params) qargs;</td><td>Declare an opaque gate</td><td>(see text)</td></tr>
<tr><td>// comment text</td><td>Comment a line of text</td><td>// oops!</td></tr>
<tr><td>U(theta,phi,lambda) qubit|qreg;</td><td>Apply built-in single qubit gate(s)</td><td>U(pi/2,2*pi/3,0) q[0];</td></tr>
<tr><td>CX qubit|qreg,qubit|qreg;</td><td>Apply built-in CNOT gate(s)</td><td>CX q[0],q[1];</td></tr>
<tr><td>measure qubit|qreg -> bit|creg;</td><td>Make measurement(s) in `z` basis</td><td>measure q -> c;</td></tr>
<tr><td>reset qubit|qreg;</td><td>Prepare qubit(s) in `|0\rangle`</td><td>reset q[0];</td></tr>
<tr><td>gatename(params) qargs;</td><td>Apply a user-defined unitary gate</td><td>crz(pi/2) q[1],q[0];</td></tr>
<tr><td>if(creg==int) qop;</td><td>Conditionally apply quantum operation</td><td>if(c==5) CX q[0],q[1];</td></tr>
<tr><td>barrier qargs;</td><td>Prevent transformations across this source line</td><td>barrier q[0],q[1];</td></tr>
</table>

Figure 31. Open QASM language statements (version 2.0). Figure from [355, 383].

To provide an example of how the main the dataset is built (to be later separated into train/test/dev datasets) using the *n -> n+1* rule and the statements mapping, is shown an Open QASM implementation of *Deutsch–Jozsa algorithm* [384] using two qubits (as it appears in [383]).

```
OPENQASM 2.0;
include "qelib1.inc";
qreg q[5];
creg c[5];
x q[4];
h q[3];
h q[4];
cx q[3],q[4];
h q[3];
measure q[3] -> c[3];
```

After mapping the code to the simpler keys, the following code is obtained (the line feed between the different lines has been removed, making the code a single line):

```
O; I; Q5; C5; X4; H3; H4; CX34; H3; M33;
```

Later, the dataset for the mapped code is augmented in different lines, where each line follows the *n -> n+1* rule, and a line below another represents *n = n-1*. So, using the mapped code, the following logical code sequences are got (Table 29).

Table 29. All the possible logical code sequences for the mapped *Deutsch–Jozsa algorithm* following the augmentation rules.

Input sequence (given one)	Output sequence (target)
O; I; Q5; C5; X4; H3; H4; CX34; H3;	O; I; Q5; C5; X4; H3; H4; CX34; H3; M33;
O; I; Q5; C5; X4; H3; H4; CX34;	O; I; Q5; C5; X4; H3; H4; CX34; H3;
O; I; Q5; C5; X4; H3; H4;	O; I; Q5; C5; X4; H3; H4; CX34;
O; I; Q5; C5; X4; H3;	O; I; Q5; C5; X4; H3; H4;
O; I; Q5; C5; X4;	O; I; Q5; C5; X4; H3;
O; I; Q5; C5;	O; I; Q5; C5; X4;
O; I; Q5;	O; I; Q5; C5;
O; I;	O; I; Q5;
O;	O; I;

Using this technique, it is possible to train the *seq2seq* network to make it able to consider what are the logical steps concerning Open QASM statements (and their arguments) used to code a quantum program, in a similar mode to the *seq2seq* training to translate between different languages..

To test this approach, some small datasets were built (typically including several thousands of code sequences) to train the *seq2seq* networks and validate the main idea. In the case of these first tests, mapping algorithm was used to discard the arguments that accompanied the statements in the original code. It could be useful to validate if the *seq2seq* networks are learning the code grammar properly (which could be easier to discover considering the small size of the initial datasets tested). For example, in a typical quantum program, after defining a register of qubits it is common to define a register of classical bits. In the tests, it does not matter if the number of qubits is the same as the number of bits or other details. The intention is to validate if the neural network has learned that a declaration of a register of classical bits should follow a declaration of a register of qubits. The initial results with the training, using the simple mapping, show that the *seq2seq* network achieves a prediction accuracy of 88-92% by comparing the training and test datasets (using the default decoder predictor of just one statement more in the target sequence per each sequence input). In the case of this research, also was introduced quantitative testing of the results provided by the dataset. In this qualitative testing, was tried to verify if the results provided by the default predictor and the predictor based on beam search are comprehensive and logical in the context of quantum programming. This qualitative testing is currently ongoing, and it involves some IBM Q team members with expertise in quantum coding who will provide their impressions of the results achieved.

Also, as part of the initial results, have been designed two main approaches to deploy the intelligent system in real contexts and integrate the assistant in the IBM Q products. These two approaches are the following:

Build an API to serve the results of the predictions and submit new codes to continue training the *seq2seq* network. In this case, a *deep learning as a service* approach is followed. The neural network and the different features are available in the cloud by using a REST API (currently implemented using Flask). Using this solution, the different products that will include the intelligent code recommender solution only need to submit the contents of the quantum program being written by the user to obtain what would be the next statement to use.

Embed the trained model within the products itself. Considering that the training model is saved into a *.pt* file of about 3MB, it can be embedded as a file within a website or a program and use it by employing [385] or introducing the code related to the predictions.

As stated, of these two approaches, the researchers have built only the deployment with the REST API. It will be used in the full validation of the system. In the other case of the embedded code, implementing it will depend on the final success of the full tests.

3.4.5 Discussion

Considering that this experience is a work in progress, and the results achieved in the beginning, the following points need to be addressed to discuss in a deeply mode the validation and final possibility of using this system:

1. Finalize the internal validation of the simpler version of the Open QASM recommender. This is fundamental to understand whether the predictions fulfill the expectations or if it is needed to redefine the intelligent system (tuning the seq2seq network, changing the neural network used, etc.).
2. Build a full dataset to train the neural network. IBM Q has over 2 million records of code executions. The plan is to build an augmented dataset using the *n -> n+1* rule and include all the code introduced by users to train the future neural networks.
3. Validate internally at IBM Q the final results of these full pieces of training. At this phase, the author will validate the results raised by the neural networks using the simple mapping as well as using a complex mapping (which maps all the statements and their arguments).
4. If all those validations and pieces of training are successful, the code recommender should be deployed using the API REST. This API would

be used by the other products and explore the integration of the code developed into other IBM Q products.

5. Using the deployed version in real products, measure different metrics of real user experiences to assess the effect of the intelligent system in the programmers' performance. In this case, will be measured (prospectively) the ratio between the code offered by the system and those statements employed finally by the users, the time they need to complete their programs using the helper, and the users' opinion through questionnaires, focus groups or other assessment tools typically used in the Human-Computer Interaction research area.

Finally, the intelligent system should regularly be re-trained to adapt its knowledge to the new code produced by users. In a new context like quantum programming, it is possible that the number of users coding as well as the complexity of the code produced will grow. The system will need to be ready to fulfill the expectations of different users with different level of coding skills in quantum computing.

4 Overall discussion

This chapter is devoted to the discussion of the results obtained during the development of the experiments and literature review performed during the thesis. It also attempts to respond to the different research questions posed in section 1.2. The structure of this chapter is the following: first, each sub-question will be presented accompanied by a reflection on how the outcomes from the research process followed contributed to respond to that sub-question. Later the primary questions will be addressed. Finally, this chapter closes with further reflections on the research carried out in this thesis. For clarification, the discussion presented in this chapter complements the discussion presented in each case study. It is important to remark that the comments below will be oriented to general aspects of the whole thesis rather than comment on specific results and other issues that have been previously commented.

4.1 Answering the research sub-questions

First question: *What kind of software artifacts does a system require to respond to the users' needs or to be adaptive to users?*

According to the experience gained during the experiments and the research in the scenarios (subsections 3.1.4, 3.2.4, 3.3.4, and 3.4.4), as well as in the literature review, any system should have the following software artifacts (components):

1. A system to gather information from the users and their interactions.
2. A system to analyze the information retrieved.
3. A system to provide feedback to different stakeholders and components based on the analysis.

These software artifacts are discussed below.

Considering the systems to gather information from the users and their interactions, the state-of-the-art exposed by the literature reveals that the universal rule in a system is to have specific components or tools acquiring and keeping information from different sources, and probably using automated techniques to do that. On the contrary, in the case of the information related to the users, to the interaction or to other HCI-related aspects, it is remarkable that not much effort has been devoted to create specific information gathering tools to do that, or at least, many environments are not designed to include this kind of systems together with the other software components. It is true that nowadays many software include tools that collect information from users for different purposes, but in most cases they are not part of the system (coming from third-party vendors) or are used to gather data and analyze users' behavior or interaction patterns for other systems rather than the environment where the users are. Some examples of this situation are presented next. The examples are related to web pages due to its simplicity and widespread application. Probably the first example of systems used to gather information from users that comes to mind is Google Analytics (GA) (https://goo.gl/XkH5u2). This system captures data from users' as they navigate in a website (which includes a Javascript snippet provided by GA) that later can be used to show reports about users, consumption patterns in the website, acquisition channels, etc. Many websites employ only Google Analytics as a tool to count unique users, visits, most visited pages, etc. and consider it a standard plugin to have on any webpage. Google Analytics is not a wrong solution. The issue to highlight in this case is the culture around this kind of systems: owners and administrators do not plan properly the resources that gather information from their users. They rely on an external system all their strategy to know their users. By doing so, they rely on a system that collects fixed information and report *small* fixed parts of knowledge (despite Google Analytics and other software products are a perfect fit for several types of users). Regarding other information retrieval artifacts, a clear example are the trackers used for ad-targeting. Similarly to the previous case, many websites integrate trackers that retrieve information from users to offer ads and track people on different websites. The tracking itself is not a problem (each business or website should know what kind of adverts allows). However, it can be the subject of a cultural issue when these admins share the data about the interaction of their users to other parties without making a proper (and ethical) exploitation of this kind of data. Overall, understanding the value of a suitable data retrieval to later analyze the information and extract knowledge is crucial for a technological project or company right now.

As for how to implement these environments, there are as many options as technologies available in the market. In the different experiments previously presented, very different strategies have been used: in three out of the four experiments, there were no available systems retrieving information about users regarding how they use the system. In those cases, it was necessary to approach this void developing software such as crawlers (experiments related to MOOCs, section 3.2), servers with APIs that receive streams of information from a closed environment (experiments related to virtual world, section 3.1) or a set of software components coupled to a bigger environment to retrieve the interaction from users in real time to be later analyzed (experiments related to large forms, section 3.3). The only case where the information about users was previously gathered from the system was the case of providing feedback to quantum computing programmers (section 3.4). As a result, in the latter case, the interaction retrieved is very different from the other cases: it is the code developed by the programmers. There are many other ways of implementing a retrieval system: many other modes for transmitting information (queues, data streaming platforms), other modes to store the information (data lakes, OLAP cubes, Hadoop-based systems, etc.), as well as other kinds of connectors depending on the technologies employed by the environment used by users. In this sense, there is no recommendation further than *have a set of tools that collect information about users' interaction empower the system and administrators to know users and use the knowledge to change the environment to face up users' needs and behaviors.*

The second software artifact needed is a system to analyze the information retrieved. This artifact, from a theoretical point of view, is a tool that extracts knowledge from the information gathered, following the principles of KDD (as presented in the Introduction). From a practical (and a software engineering) point of view, the structure of this system can be based on software interfaces to interconnect with other systems (APIs, messages, and so on, as it has been the case with the experiments performed) and a core component that will implement the analysis methods and routines. This core component or, to be more precise, these methods for the analysis implemented in the core, will be dependent on the problem to solve. In the case of this thesis, these methods have been different: in two cases, virtual worlds (section 3.1) and MOOCs (section 3.2) simple statistical tests to enable descriptive analysis have been implemented; in the two remaining cases, different AI methods were used. In the case of the large web forms (section 3.3), ML was used to detect the best version of the questionnaire for each kind of users and to cluster the users based on their similarities; and in the case of programming in quantum computing (section 3.4) a DL approach has been used to understand the code sequences employed by the users.

Despite the evident difference between the different systems to analyze the users' interaction, all of them have proven to be valid in the case studies. It is true that the methods based on AI can enable new ways of analyzing and tackling problems, but by no means they are the unique approach to take into account to solve the challenges raised in supporting, analyzing or enhancing human-computer interaction. Also, the AI approaches will be discussed below.

The third software artifact proposed is a system to provide feedback to different stakeholders and components based on the analysis. This artifact is especially sensitive. It could require more work than the other two artifacts. There are many ways of giving feedback to users, stakeholders or other software components. The feedback to be provided will change depending on the nature of the recipient of the feedback. For that reason, this software artifact can require the expertise of people related to software engineering, data analysis, UX, visual design, psychology, etc. In the thesis four different ways of providing feedback to different users have been explored:

1. Textual (messages) feedback to users when they request it.
2. Automated textual feedback (i.e., the request by users is not required).
3. Visual feedback to users: data visualization.
4. Feedback based on changing the visual layout of the system.

Concerning the textual feedback, it has been employed on three scenarios, in those related to: virtual worlds, MOOCs and supporting quantum computing programmers. On-demand textual feedback has been used in the virtual worlds scenario. In this case, as seen in Figure 10, if the users clicked on a specific object within the Usalpharma laboratory, it opened a dialog box informing the students' performance in checking the different facilities of the laboratory that were relevant for the audit. The other case that employed feedback using text was elated to the MOOCs experiments. In these experiments, MOOC teachers and administrators were given datasets and text files with information about the users' interaction in the MOOC and the social network to be later analyzed. In this case, the feedback was in raw mode, the tools (crawlers and software architecture) provide the raw data to the different stakeholders involved, making them responsible for discovering the knowledge. Moreover, the automated textual feedback has been used in the experiments related to supporting programmers in QC. Feedback is automatically provided to the users since the system evaluates the initial input from the user and propose next pieces of code or statements to be used according to what the *seq2seq* neural networks have learned.

Moreover, the visual feedback has been used in the virtual worlds scenario as a way to give information to the teachers and staff responsible of managing the virtual facilities to inform them about the students' interaction and virtual world usage. In this case, the visual feedback was provided using graphical visualizations of the usage and interaction data and composing visual reports in a website.

The last method to provide feedback used during the thesis consisted in changing the visual layout of the system. This method has been used in the case study of improving UX in large forms. This kind of feedback could may not be perceived by the users: for example, in the experiments presented in section 3.3, some users do not use more than one layout because they have been profiled previously and are redirected to the layout that is a best fit for them. Also, there are other users that experiment two different layouts. In these group of users that use more than one visual interface, the change has been demonstrated to be effective, but there are no evidence that users were aware of the change.

These three systems —quasi-theoretical and coarse-grained— proposed to answer the question are simple artifacts proposed as the minimal software components needed to gather information, process it and return value to the users of a system. In each different environment or software project, these minimal components should be decomposed into other more specific or specialized ones. Also, they can be complemented with third-party systems or subsystems to delegate some part of the tasks (decisions, algorithms, information retrieval, etc.) and improve the overall results achieved. According to the literature, these components are present in part of the works analyzed (with different names but with the same purpose), but they lack standardization. Moreover, a lot of internal systems are found in the industry that track users or analyze their interaction using similar schemas to those proposed here, but not many of them are used to improve the user experience; they are mostly used to enable advertising. Nevertheless, some successful companies use this kind of approach to improve the UX of a system. Probably the best example of using this approach is Netflix, which has revealed recently that it uses similar methods to those presented in this thesis [386-388]. It seems reasonable to think that more companies are using these approaches, but most of them do not make public their business strategies.

Concerning the question *what features should have these kinds of systems?* The answer is based on the previous one and in the outcomes of the experiments run. From the experience gained after implementing the different components described previously, the most relevant features that should be included are:

1. Transparent data collection. The system should collect information about users without disturbing them.
2. Ethical data collection and exploitation. The system must provide adequate methods to deal with data in a respectful mode regarding the users and the different stakeholders involved in a system. The principles to deal with data ethically have been exposed in the literature [389-391] and can be summarized:
    - 2.1. Ownership: the owner of the data should be the individual (in this case, the user).
    - 2.2. Transaction transparency: the design of algorithms that use data of individuals should be available transparently.
    - 2.3. Consent: users should be informed about who control their data, the purpose of the data analysis and they must grant consent to the conditions expressed.
    - 2.4. Privacy: privacy should be protected even in data transactions (and not only in the storage of information).
    - 2.5. Currency: the users should be aware of any financial transaction resulting from the use of the data.

2.6. Openness: the aggregate datasets (anonymized) should be freely available.

Sometimes, depending on the project and its legal issues, all of these principles might not be easy to apply. Nevertheless, all of them should be included in the system during the design phase if it is possible.

Another kind of ethical treatment of the data is to avoid using the knowledge extracted from the data to enable dark patterns [392-394] in the user interfaces, designed to trick the user.

3. A well-designed data environment. The data environment should be capable to manage the data efficiently considering the needs of the data to handle and the expected outcomes from the information management. The design should cover the data transmission inside and outside the system, the data collectors, the analysis components, and all the other issues related to a data-driven system.

   This recommendation also includes the system that analyzes the data. Regarding the data analysis, the goals should be defined previously, the analysis methods should fit in the data and the expected outcomes, and it should cover the features commented in these lines.

4. Proper effectors to deal appropriately with HCI challenges. To take advantage of the knowledge gained from the information gathered, the system should have the proper elements to use the data. An effector, in this case, can be defined as a software element that produces actions on other software component or in the users. For that reason, the knowledge generated from the information should take effect on other parts of the system, directly on the users or in both at the same time. Examples of effectors are the API that returns suggested code sequences in the case study of programming in quantum computing, the objects that provide feedback to the users inside the virtual world, etc.

Having these features does not ensure an optimal quality of the environment that faces the HCI challenges, but introduces a suitable scaffold to build a better system that responds to the actual challenges.

On the question *what kind of software behaviors related to those features should be the common ones?* First of all, the meaning of behavior needs to be clarified. In the case of this explanation, behavior will be understood as the way in which software performs actions, and is not used in the sense expressed in computer science theories such as Behavior-Driven Development [395, 396], despite a similar vocabulary may be used in some cases.

In general, and related to the effectors previously described, the system should be *reactive*. According to the "The Reactive Manifesto" [397], a reactive system is:

1. Responsive: The system responds on time if possible. Responsiveness is the cornerstone of usability and utility, but more than that, responsiveness means that problems may be detected quickly and dealt with effectively.
2. Resilient: The system stays responsive in the face of failure. This applies not only to highly-available, mission-critical systems —any system that is not resilient will be unresponsive after a failure—.
3. Elastic: The system stays responsive under a varying workload. Reactive Systems can react to changes in the input rate by increasing or decreasing the resources allocated to service these inputs.
4. Message-driven: Reactive Systems rely on asynchronous message-passing to establish a boundary between components that ensure loose coupling, isolation and location transparency. This boundary also provides the means to delegate failures as messages.

Figure 32 depicts the characteristics of a reactive system and the relations between them.

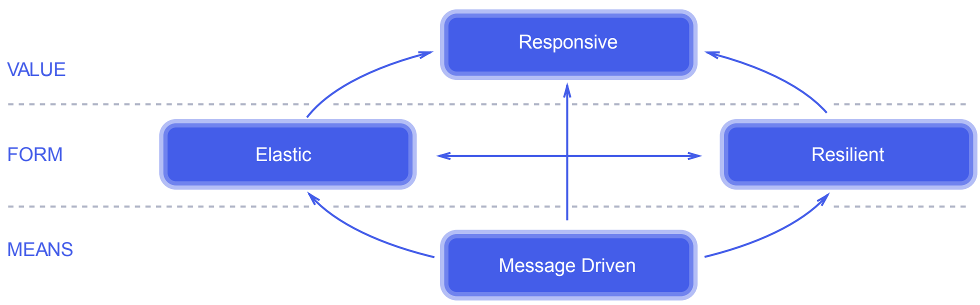


Figure 32. Overview of the traits of Reactive systems. Source: [397].

As discussed in the thesis and found in the literature, the task of retrieving interaction from users is not straightforward [43, 91, 177, 179, 398, 399]. For that reason, the philosophy of reactive systems seems a perfect fit as a model to describe/design the behavior of a software environment that deals with HCI issues [400]. In the case of retrieving interaction, the data streams coming from an interactive system can be huge (depending on the moment and the concurrency of users in the virtual platforms), and the system needs to be efficient resolving the task. Any error on transmission, any failure managing data can trigger a domino effect ending up losing more interactions and ruin the data retrieval. Also, if the feedback must be given to the users on real-time (like with the feedback provided by the Usalpharma's virtual objects), the system must respond timely.

During the thesis and the different experiments and experiences performed, these principles and behavior guides have not been applied consistently. This scarcity has sometimes driven to different issues with users, data retrieval and the quality of the feedback provided. The reader can consider the need for a reactive system as lesson learned during the development of the case studies.

The next question to answer is *what kind of strategies should these systems include to provide valuable feedback to the users?* As seen in the different experiments presented, the feedback has been provided to the users in several ways:

1. Gauzy feedback. This feedback is provided to the users in a transparent mode. The user can notice that the system is responding to their actions and be aware of the different layout changes, messages or another kind of feedback received.
2. Stealth feedback. This feedback is given to the users without the need for them to notice it. This is the case of profiling users and offering them the version of web questionnaire that is a better fit for their characteristics.

The reader may think that stealth feedback collides with the classical feedback principle definition: "[...] is related to the visibility principle and refers to sending back information from the system to the users so that they know what effect their actions have had" [325]. In the author's opinion, they are not incompatible. Stealth feedback offers some benefits to the user similar to gauzy feedback, but acts within a different timing. This timing considers ways to improve the system whether the user is connected or not, so the feedback cannot be anymore only a question of responding in real time to the users, but also to learn from the users in order to improve the experience the next time they interact. This extension of the timing where the feedback is given also agrees with the considerations on feedback timing present in the literature [56, 325].

During the thesis, the feedback has been provided under two circumstances: when users request it as well as when users do not request it. The case of the feedback requested by users has been tested in the two initial study cases: the Usalpharma virtual world (section 3.1) and the MOOCs (section 3.2). Likewise, the case of feedback not requested by users has been tested in the cases of web forms (section 3.3) and in aiding programmers (section 3.4).

In the first three case studies (all except the help to QC programming) both strategies of giving feedback have proved to be positive. All of them have raised positive results. At this point, it is possible to enquire what type of feedback could considered to be the best? According to the literature, most papers include evident (gauzy) feedback [83, 132, 148, 162]. However, in recent papers stealth feedback is gaining traction [328, 387, 401, 402] but it is not mainstream as for now. So the previous question has not an straight answer: both of them have benefits and can have different goals. In the case of gauzy feedback, it is clear that it can improve the user experience while using a system, reinforcing positive feelings towards the interaction experienced, etc. In the case of stealth feedback, it can be used to improve the contents, layout and other issues that will be experienced by the users. From the current knowledge about them, the author foresees that they can be applied in different moments of the software life cycle and of the users' interaction. Overall, it is advisable to further research the limitations of type of feedback, which could be an excellent future line to work.

The last sub-questions to answer are: *could an intelligent system be capable of improving the UX in a significant manner? Would it be adequate to include intelligent features in a system to pursue such UX improvement?*

To respond to these questions, two experiments have been developed during the thesis that include artificial-intelligence-related techniques (scenarios three and four). The approach followed when applying the AI-techniques was not to create algorithms, machines or software entities that can think by themselves [403-405], but to employ logical stakeholders able to learn from given information and propose some solution to the problem posed. During the last part of the thesis (years 2017-2018), these solutions have been applied to different problems and different methods (supervised learning, unsupervised learning, deep learning) have been used with promising results. As seen in the current literature more authors are using AI-techniques to affect users, their experience in an interactive environment or to analyze the interaction [386, 406]. According to the literature review performed during the thesis, the use of AI-related techniques have not been the typical approaches in the field of HCI, despite some of them propose it [56], so these new papers (as well as the ones written in this thesis [67]) can be considered to be opening a new era by bridging both knowledge areas together. In general, according to the authors in the literature and the experience gained in the experiments performed, an intelligent system can be a good option to work in the knowledge discovery part of the HCI-KDD process. The inherent features of intelligent systems that employ AI-related techniques or methods can help to discover hidden patterns and relations in the data from users' interaction and enable new knowledge to improve the software. So, the answer to the first sub-question is: yes, taking into account the experience gained and the results achieved during the thesis, an intelligent system can be capable of improving the UX of a system in a significant manner. The response is positive even considering that in the second experiment (the one related to quantum computing) it has not being possible to conduct extensive tests with real users, but the first impressions given by the experts consulted permit to be optimistic in this regard.

About the adequacy of applying intelligent features to a system to pursue UX improvements, the answer is not as clear as in the previous case. The adequacy depends on the goals pursued, the methods utilized and the data available. In other words, depending on the AI methods used, the results can be explainable or not, and the process followed by the algorithms can be interpretable or not (black-box models), etc. Even, the results achieved by AI techniques can be valid observations or can be the result of an overfitting phenomenon. Moreover, if the data available are not enough, many AI methods will not give adequate results. Also, in general, the knowledge achieved using this kind of techniques do not express causality [407]. So, considering these handicaps, if the goal pursued is related to obtaining an explainable or interpretable knowledge by humans, some AI-related methods cannot be recommended. This is the case with the experiments done using the deep learning approach (section 3.4). Besides, using white-box models like those employed in the third scenario (section 3.3) enable the interpretability of the results and allow researchers to track the different data transformations and intermediate results.

So, the answer is halfway. The adequacy of these methods depends exclusively on the goals pursued, the methods employed and the data available. The AI-related methods can extract knowledge from most kind of (mainly big) datasets, but it is the responsibility of the researchers to use the proper methods for the expected outcomes.

4.2 Answering the primary research questions

The final answers are related to the central research questions posed in section 1.2: *is it beneficial to follow a data-driven/KDD approach in a system to support, analyze or improve the users' interaction and experience and tackle the HCI-related challenges to these aspects that the interactive systems present? How a software environment should evolve to respond to the users' needs and improve or support the users' interaction and experience? How is it possible to do that in a more automated way?* At this point of the research, a partial answer can be obvious to the reader: the software environments should embrace KDD principles and methodologies to respond properly to the users' needs and improve or support the HCI processes, tackling the typical issues presented in systems that involve users' interaction.

After researching using four different scenarios, performing several experiments, and working to improve, support or analyze the human-computer interaction in several different systems involving more than 6000 users, the most valuable outcome for the author is that KDD (and data-driven) approaches worked in this type of contexts and helped to face the challenges proposed.

To embrace the data-driven or KDD approach, software environments need to extend their software architectures to cover also the issues related to Human-Computer Interaction. In this section some solutions have been presented regarding the minimal software artifacts to include, their behavior, the features desired in the extended software architecture, or the different kinds of feedback to reinforce the users' interaction and experience. Since many years ago, the HCI community has been seeking for new ways of automatizing the evaluation of users' interaction, adapt systems automatically to the users, retrieve information to extract knowledge, etc. [408]. The work done in this thesis does not solve the problem but shows some exciting approaches to deal with the problem from the software perspective to influence what users feel and perceive.

It is true that the outcomes resulting from the performed research are not exportable to most kinds of systems since the peculiarities in each system introduce many difficulties to solve this kind of issues in this context (at least concerning automation, metrics, and adaptation to users). HCI is too broad and includes too many areas of knowledge, approaches, paradigms and so on to be feasible to propose a kind of "unified theory" that ultimately solve the issue. Probably, in the future, this issue will be solved better than today, and the solutions will not be as ad-hoc as currently, but for now, it is the moment to propose new ways of working, new approaches and benefits from other areas of research that can be adopted by the HCI community.

4.3 Final reflections

Some final remarks can be added concerning the work carried out, that outline some issues related to the case studies that have not been discussed before.

Although previously it was mentioned the involvement of 6000 users, the final results presented in section 1 include analyses of about 1500 users of interactive systems. This is because in the case study of large web forms more than 4000 users were discarded during the cleaning stage necessary to perform the machine learning algorithms. Also, in the case of the fourth study, it must be added that it does not include data about users, but it includes around 500,000 QASM code sequences.

Regarding the interest in the results of each case study, the author wants to highlight the most relevant and the less one. For the author, the most relevant case study developed in the thesis is the work done on improving UX and enabling adaptability of large web forms using machine learning. That work, besides including more users than any other case, is where the most relevant results were achieved. These results are of several types: a traceable machine learning workflow, the application of different algorithms that could be interpretable by researchers, the improvement of users' performance, as well as to be the most valuable demonstration achieved on how AI techniques can be applied to face HCI challenges. On the contrary, the least relevant case study is that related to MOOCs. This case was planned in the beginning as an excellent opportunity to experiment with a high amount of users (as is typically the case in MOOCs and social networks) and with new kind of systems that usually need the extension of their features using new software components. In this case, the software components designed and the overall software architecture could be the most exciting results. Regarding the users, some issues reduced the interest of the possible results: First of all, the users enrolled in the MOOC course were not as many as expected (there were some MOOCs organized in the same platform that involved more than 500 users). Also, the conversations and interactions between users in the different social networks analyzed were not as meaningful as expected.

As for the remaining two case studies, the following can be commented. In the case of the experiments performed in the Usalpharma virtual facilities, the case study was the most successful in enabling users to do things and perform actions that previously were not possible. Although the research results are less attractive than those achieved in the case of the web forms and machine learning, the results regarding the software and the utility of the work performed were highly satisfactory for the different stakeholders involved (teachers, administrators, students). Lastly, regarding the work in progress project of aiding programmers in quantum computing, the project seems to be an excellent opportunity to advance in improving the human-computer interaction in a new application field. However, since there are no results involving evaluation with real users, it cannot be considered to be at the same level as the other successful case studies. Nevertheless, the author hopes to continue working on this research line and achieve the expected results.

5 Conclusions and further work

The primary outcomes of the research undertaken for this thesis were the literature review and mapping of software architectures affecting human-computer interaction and the development of different software solutions in different scenarios to analyze, support and enhance the users' interaction and experience. On the one hand, and according to the conducted literature review and mapping, many authors have been working along the years in the definition, development, and testing of software solutions based on software architectures to deal with the analysis and enhance of users' interaction with computers. Despite this effort, the literature shows that there are many applications based on ad-hoc solutions and a lack of formal proposals and standardization. In the review, a lack of intelligent software that could enable new ways of interaction analysis, feedback to users, etc. was also observed. Taking into account the achieved results, it is clear that the advance in the integration of new trends like data-driven environments, artificial intelligence or reproducible workflows for analysis and feedback of users can lead to a significant improvement in the research area. Moreover, there is a need of conducting research focused on the definition of standardized techniques and novel approaches to deal with the analysis, support, and enhancement of the users' interaction and experience using software environments. Besides this, and related to the literature, it has also been detected a significant number of publications outside the academic publications that deal with similar goals than those specified in this thesis. In this case, it was found that many companies are applying similar techniques and approaches to those presented in this manuscript in the context of users' engagement with digital products. So, this situation suggests that the academia is missing a good opportunity to establish the foundations for new ways to extract knowledge from users' interactions and experience.

On the other hand, the research and development of software solutions to analyze, support and enhance the users' interaction and experience carried out in the different case studies has been conducted following the recommendations given in the literature and considering the different needs and possible improvements detected in the reviewed literature. In this sense, in the performed empirical research, the following aspects prevailed: 1) the definition of data-driven software systems; 2) the development of clear information workflows that could help different stakeholders with the issues related to the users' interaction and experience; and 3) the employment of novel techniques to analyze the available data and provide feedback consequently the users.

Considering the results yielded after the empirical research, the main conclusion of this work is the evidence of the positive effect of embracing KDD principles to deal with the HCI challenges related to the improvement of users' experience and interaction. Using the KDD approach, systems can be built to include different software entities that would take advantage of the information generated by the users and their interaction with the software. Also, if the information workflows and goals for analysis are well defined, these software environments could automate the analysis and feedback for the users' interaction and experience. These assertions are supported by the conducted experiments, which have involved more than 6000 users. All the case studies in this research have followed an approach based on knowledge discovery from data, which fulfills the proposed research objectives. Despite some of the software solutions presented in the manuscript are not fully transferrable outside the case studies where they were applied, they represent some shared software artifacts and techniques which could be compatible with other cases, after the proper adaptation. One remark about the transferability of the solutions presented is that some common metrics and indicators could be used in different scenarios, but it is highly complicated to define a closed set of them that could be applied to any system. Human-Computer Interaction is a challenging field that contains many sub-fields and paradigms, and each one of them presents different needs and problems. In this case, despite the difficulty, in this thesis some common generic software artifacts have been applied in most cases, which could be used in almost any scenario. In this sense, the employment of replicable techniques, the public release of analysis workflows or the provision of open data for research can help the researchers in HCI. These resources can be used to expand the area and achieve further significant results like the creation of data pipelines to automate the analysis and improvement of users' experience or to develop report systems on users' interaction to understand the users and their needs and fulfill their expectations.

On the need of extracting knowledge, during the experimental part of this thesis it has been demonstrated that many different kinds of stakeholders (humans and software ones) can take advantage of the knowledge raised by the human-computer-interaction related information. Such a deeper understanding could help these stakeholders to evaluate users' interaction, to detect users' behavior, to extract performance patterns, to engage users in different tasks, or to help them to solve a difficult task.

Concerning the different techniques used to analyze data and extract knowledge, those that present a major novelty and the most promising results are the AI-related techniques and its subareas (machine learning, deep learning). The analysis techniques to be applied in each case depend on the problem to solve, but using artificial intelligence and approaches like supervised learning, unsupervised learning or natural language processing can be adequate to automate part of the knowledge extraction tasks and build workflows to tackle some common challenges in analyzing, support and enhance the users' experience and interaction. This conclusion is consequent with the workflow designed for the third research scenario, which was employed in two different problems achieving significant results on both.

As a coda to the conclusions, and strengthening some of the ideas available in the literature, the analysis of data related to interactive systems can lead to improving them significantly. By establishing the adequate methods to do that, developers, designers, and managers of interactive software can better understand the users' needs and desires and fulfill them. The definition of software subsystems within interactive scenarios responsible for analyzing and improving users' interaction and experience since the beginning of a digital product should be mandatory. These subsystems can lead to develop earlier positive user's feelings, enforce better users' perception and achieve different levels of trustiness, reliability or engagement in the users. This idea fits perfectly also in research approaches like the technological ecosystems, where many different components collaborate to pursue global goals. The technological ecosystems (and in their information systems) present many challenges related to the interoperability between the components, the complexity in data flows or the timely and adequate response to the users (which are also considered a part of the ecosystems). The results presented in this thesis on extended (and more flexible) software architectures and the application of intelligent approaches could open new possibilities for tackling these challenges.

5.1 Further work

There are several opportunities for further research based on the research presented in this thesis. First of all, there are many topics to be investigated related to the relationship between artificial intelligence and the human-computer interaction. The first topic could be the analysis of other different applications of AI techniques to evaluate the users' interaction and experience and try to detect which of them are suitable to automate the improvement of users' experience. The second one could be a study on what kind of AI techniques could help researchers and software engineers to create adaptive HCI-related systems.

Outside AI, there are opportunities to expand this research in the definition of a catalog of white-box algorithms and analytics workflows to extract knowledge from data and to explain the results in the area of analyzing users' interaction. Also, it would be very valuable to define a catalog of relevant metrics to be included in any system that aims to analyze, support or improve HCI processes.

Within the software engineering knowledge area, the standardization (through software patterns, meta-models, etc.) of the ideas related to software architectures is a challenging research, in order to create a general framework and contribute to the development of new HCI-centered systems. Also, it would be very significant to develop software frameworks that could be pluggable to other software environments or technological ecosystems to extend them using the results of the standardization outlined above.

Similarly, in the HCI knowledge area, a further research line is to continue the work performed in the case study of large forms to analyze the effect of each proposed change in the form on the users' performance, trust, and experience.

Another research line that could be relevant to continue the work done in this thesis is to propose better methodological models to support the experimentation with users in systems that evolve to support or improve HCI processes.

Finally, and keeping in mind how the ideas and results presented could benefit other research fields, it could be extremely interesting to apply the different ideas and outcomes from this thesis, for example, to improve current eLearning systems, in coordination with research areas like learning analytics, PLEs, etc. The experimentation in real systems with different types of users and stakeholders can be the source of opportunities to discover new possibilities of improvement.

5.2 Outcomes from this thesis

This final subsection outlines the different merits achieved by the candidate during the thesis, concerning the publications in different channels, the software released, the intellectual property records recognized, the awards and grants received during the thesis and the pre-doctoral research stay performed.

5.2.1 Publications

Journal papers

1. [354] F. J. García-Peñalvo, J. Cruz-Benito, M. Martín-González, A. Vázquez-Ingelmo, J. C. Sánchez-Prieto and R. Therón, "Proposing a Machine Learning Approach to Analyze and Predict Employment and its Factors," *International Journal of Interactive Multimedia and Artificial Intelligence,* vol. In Press, 2018. doi: 10.9781/ijimai.2018.02.002. (ESCI)
2. [67] J. Cruz-Benito, A. Vázquez-Ingelmo, J. C. Sánchez-Prieto, R. Therón, F. J. García-Peñalvo and M. Martín-González, "Enabling Adaptability in Web Forms Based on User Characteristics Detection Through A/B Testing and Machine Learning," IEEE Access, vol. 6, pp. 2251-2265, 2018. doi: 10.1109/ACCESS.2017.2782678. (JCR SCI - COMPUTER SCIENCE, INFORMATION SYSTEMS - Q1 (24 de 148);

ENGINEERING, ELECTRICAL & ELECTRONIC - Q1 (48 de 260); TELECOMMUNICATIONS – Q1 (19 de 87) - IF 3.557)

3. [297] J. Cruz-Benito, O. Borras-Gene, F. García-Peñalvo, A. Blanco and R. Theron, "Learning communities in social networks and their relationship with the MOOCs," *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje,* vol. PP, no. 99, pp. 1-1, 2017. doi: 10.1109/RITA.2017.2655218. (SJR 0.206 – ENGINEERING (MISCELLANEOUS) – Q3; EDUCATION – Q3; E-LEARNING – Q3)
4. [175] J. Cruz-Benito, C. Maderuelo, F. J. García-Peñalvo, R. Therón, J. S. Pérez-Blanco, H. Zazo, A. Martín-Suárez, "Usalpharma: A Software Architecture to Support Learning in Virtual Worlds," *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje,* vol. 11, no. 3, pp. 194-204, 2016. doi: 10.1109/RITA.2016.2589719. (SJR 0.206 – ENGINEERING (MISCELLANEOUS) – Q3; EDUCATION – Q3; E-LEARNING – Q3)
5. [318] J. Cruz-Benito, O. Borrás-Gené, F. J. García-Peñalvo, Á. Fidalgo Blanco and R. Therón, "Comunidades de Aprendizaje en Redes Sociales y su Relación con los MOOC," *VAEP RITA,* vol. 4, no. 2, pp. 87-99, 2016.
6. [256] J. Cruz-Benito, R. Therón, F. J. García-Peñalvo and E. Pizarro Lucas, "Discovering usage behaviors and engagement in an Educational Virtual World," *Computers in Human Behavior,* vol. 47, no. 0, pp. 18-25, 6// 2015. doi: 10.1016/j.chb.2014.11.028. (JCR SSCI – PSYCHOLOGY, MULTIDISCIPLINARY – Q1 (21 de 129); PSYCHOLOGY, EXPERIMENTAL – Q1 (20 de 85) – IF 2.880)
7. [251] J. Cruz-Benito, C. Maderuelo, F. J. García-Peñalvo, R. Therón, J. S. Pérez-Blanco, H. Zazo, A. Martín-Suárez, "Usalpharma: Una arquitectura software al servicio del aprendizaje en Mundos Virtuales," *VAEP RITA,* vol. 3, no. 3, pp. 148-159, 2015.
8. [178] C. Maderuelo, A. Martin-Suarez, J. S. Pérez-Blanco, H. Zazo, J. Cruz-Benito and A. Domínguez-Gil, "Facility-based inspection training in a virtual 3D laboratory," *Accreditation and Quality Assurance,* vol. 19, no. 5, pp. 403-409, 2014/10/01 2014. doi: 10.1007/s00769-014-1065-4.

(JCR SCI – CHEMISTRY, ANALYTICAL – Q4 (60 de 74); INSTRUMENTS & INSTRUMENTATION – Q3 (40 de 56) – IF 0.966)

9. [188] F. J. García-Peñalvo, J. Cruz Benito and R. Therón Sánchez, "Visualización y Análisis de Datos en Mundos Virtuales Educativos: Comprendiendo la interacción de los usuarios en los entornos 3D," 2014.
10. [179] F. J. García-Peñalvo, J. Cruz-Benito, C. Maderuelo, J. S. Pérez-Blanco and A. Martín-Suárez, "Usalpharma: A Cloud-Based Architecture to Support Quality Assurance Training Processes in Health Area Using Virtual Worlds," The Scientific World Journal, vol. 2014, art. 659364, p. 10, 2014. doi: 10.1155/2014/659364. (SJR 0.392 – BIOCHEMISTRY, GENETICS AND MOLECULAR BIOLOGY (MISCELLANEOUS) – Q2; ENVIRONMENTAL SCIENCE (MISCELLANEOUS) – Q2; MEDICINE (MISCELLANEOUS) – Q2)
11. [202] A. Martin-Suarez *et al.*, "Scientific Knowledge Transfer Training Through a Virtual World," *Journal of Information Technology Research (JITR),* vol. 7, no. 2, pp. 24-35, 2014. doi: 10.4018/jitr.2014040103. (SJR 0.119 – COMPUTER SCIENCE (MISCELLANEOUS) – Q4)

Conference papers

1. [363] J. Cruz-Benito, I. Faro, F. Martín-Fernández, R. Therón and F. J. García-Peñalvo, "A Deep-Learning-based proposal to aid users in Quantum Computing programming," presented in HCI International 2018, Las Vegas, USA, 2018. doi: 10.1007/978-3-319-91152-6_32.
2. [321] J. Cruz-Benito, J. C. Sánchez-Prieto, A. Vázquez-Ingelmo, R. Therón, F. J. García-Peñalvo and M. Martín-González, "How different versions of layout and complexity of web forms affect users after they start it? A pilot experience," in *World Conference on Information Systems and Technologies*, 2018, pp. 971-979: Springer.
3. [409] A. Vázquez-Ingelmo, J. Cruz-Benito and F. J. García-Peñalvo, "Improving the OEEU's data-driven technological ecosystem's interoperability with GraphQL," presented in Fifth International Conference Technological Ecosystems for Enhancing Multiculturality 2017 (TEEM'17), Cádiz, Spain, October 18-20, 2017, 2017.

4. [410] R. Therón, J. Cruz-Benito, F. García-Sánchez, R. Santamaría and F. García-Peñalvo, "Innovación en la enseñanza de la Interacción Persona-Ordenador: interfaces imaginadas, ciencia-ficción y trabajo con usuarios reales," presented in IV Congreso Internacional sobre Innovación, Aprendizaje y Competitividad, CINAIC 2017, Zaragoza, Spain, 2017.
5. [411] J. Gómez Isla, F. García-Sánchez, J. Cruz-Benito and C. González García, "Procesos colaborativos de crítica y reflexión para la coevaluación de proyectos artísticos de alumnos de Bellas Artes mediante el uso de las tecnologías móviles," presented in IV Congreso Internacional sobre Innovación, Aprendizaje y Competitividad, CINAIC 2017, Zaragoza, Spain, 2017.
6. [322] J. Cruz-Benito *et al.*, "Improving Success/Completion Ratio in Large Surveys: A Proposal Based on Usability and Engagement," in *Learning and Collaboration Technologies. Technology in Education: 4th International Conference, LCT 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part II*, P. Zaphiris and A. Ioannou, Eds. pp. 352-370, Cham: Springer International Publishing, 2017. doi: 10.1007/978-3-319-58515-4_28.
7. [42] J. Cruz-Benito, R. Therón and F. J. García-Peñalvo, "Software Architectures Supporting Human-Computer Interaction Analysis: A Literature Review," in *Learning and Collaboration Technologies: Third International Conference, LCT 2016, Held as Part of HCI International 2016, Toronto, ON, Canada, July 17-22, 2016, Proceedings*, P. Zaphiris and A. Ioannou, Eds. pp. 125-136, Cham: Springer International Publishing, 2016. doi: 10.1007/978-3-319-39483-1_12.
8. [412] J. Cruz-Benito, O. Borrás-Gené, F. J. García-Peñalvo, Á. Fidalgo Blanco and R. Therón, "Detección de aprendizaje no formal e informal en Comunidades de Aprendizaje soportadas por Redes Sociales en el contexto de un MOOC Cooperativo," in *Actas del XVII Simposio Internacional de Informática Educativa (SIIE'15) (Setúbal, Portugal. 25-27 November 2015)*, M. d. R. r. Rodrigues, M. n. Llamas Nistal and M.

Figueiredo, Eds. pp. 410-418, Portugal: Escola Superior de Educação do Instituto Politécnico de Setúbal, 2015.

9. [249] J. Cruz-Benito, O. Borrás-Gené, F. J. García-Peñalvo, Á. Fidalgo Blanco and R. Therón, "Detection of Non-Formal and Informal Learning in learning communities supported by social networks in the context of a Cooperative MOOC," in *Proceedings of the XVII International Symposium on Computers in Education (SIIE'15) (Setúbal, Portugal. 25-27 November 2015)*, M. d. R. r. Rodrigues, M. n. Llamas Nistal and M. Figueiredo, Eds., Portugal: IEEE, 2015.
10. [413] A. García-Holgado, J. Cruz-Benito and F. J. García-Peñalvo, "Comparative analysis of the Knowledge Management in Spanish Public Administration," presented in III Congreso Internacional sobre Aprendizaje, Innovación y Competitividad - CINAIC 2015, Madrid, 2015.
11. [414] A. García-Holgado, J. Cruz-Benito and F. J. García-Peñalvo, "Analysis of Knowledge Management Experiences in Spanish Public Administration," presented in Third International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'15), Porto, Portugal, 7-9, October, 2015.
12. [109] F. García-Sánchez, J. Cruz-Benito, R. Therón and J. Gómez-Isla, "Designing and building systems and tools to analyze visual communications on social networks," presented in Third International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'15), Porto, Portugal, 2015. doi: http://dx.doi.org/10.1145/2808580.2808629
13. [333] F. Michavila, M. Martín-González, J. M. Martínez, F. J. García-Peñalvo and J. Cruz-Benito, "Analyzing the employability and employment factors of graduate students in Spain: The OEEU Information System," presented in Third International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'15), Porto, Portugal, 2015.
14. [43] J. Cruz-Benito, O. Borrás-Gené, F. J. García-Peñalvo, Á. Fidalgo Blanco and R. Therón, "Extending MOOC ecosystems using web services and software architectures," in *Proceedings of the XVI International*

Conference on Human Computer Interaction pp. 1-7, Vilanova i la Geltrú, Spain: ACM, 2015. doi: 10.1145/2829875.2829923

15. [250] F. J. García-Peñalvo, J. Cruz-Benito, O. Borrás-Gené and Á. Fidalgo Blanco, "Evolution of the Conversation and Knowledge Acquisition in Social Networks Related to a MOOC Course," in *Learning and Collaboration Technologies*, vol. 9192, P. Zaphiris and A. Ioannou, Eds. Lecture Notes in Computer Science, pp. 470-481, Switzerland: Springer International Publishing, 2015. doi: 10.1007/978-3-319-20609-7_44.
16. [415] J. Cruz-Benito, R. Therón and F. J. García-Peñalvo, "Analytics of information flows and decision making in heterogeneous learning ecosystems," in *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality*, 2014, pp. 703-707: ACM
17. [203] J. Cruz-Benito, F. J. García Peñalvo, R. Therón, C. Maderuelo, J. S. Pérez-Blanco, H. Zazo, A. Martín-Suárez, "Using software architectures to retrieve interaction information in eLearning environments," in *Computers in Education (SIIE), 2014 International Symposium on* pp. 117-120, 2014. doi: 10.1109/SIIE.2014.7017715.
18. [177] J. Cruz-Benito, F. J. García-Peñalvo and R. Therón, "Defining generic data collectors for Learning Analytics: Facing up the heterogeneous data from heterogeneous environments," presented in International Conference on Advanced Learning Technologies (ICALT) 2014, Athens, Greece. , 2014. doi: 10.1109/ICALT.2014.108.
19. [176] J. Cruz-Benito, R. Therón, F. J. García Peñalvo, C. Maderuelo, J. S. Pérez-Blanco, H. Zazo, A. Martin-Suarez, "Monitoring and feedback of Learning Processes in Virtual Worlds through analytics architectures: A real case," in *Sistemas y Tecnologías de Información. Actas de la 9ª Conferencia Ibérica de Sistemas y Tecnologías de Información*, vol. I Artículos, Á. Rocha, D. Fonseca, E. Redondo, L. P. Reis and M. P. Cota, Eds. pp. 1126-1131, Barcelona, España, June, 18-21, 2014: AISTI (Asociación Ibérica de Sistemas y Tecnologías de Información), 2014.

20. [204] A. Martín-Suárez, J. Cruz-Benito, J. S. Pérez-Blanco, M.d.C. Gutiérrez Millán, A. Zarzuelo Castañeda, M. J. de Jesús Valle, H. Zazo Gómez, C. Maderuelo Martín, J. M. Armenteros del Olmo, J. M. Lanao, "Virtual congresses for pharmaceutical learning," presented in Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality, Salamanca, Spain, 2013. doi: 10.1145/2536536.2536549.
21. [215] J. Cruz-Benito, R. Therón, F. J. García-Peñalvo and E. Pizarro Lucas, "Analyzing users' movements in virtual worlds: discovering engagement and use patterns," presented in Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality, Salamanca, Spain, 2013. doi: 10.1145/2536536.2536622.

Book chapters

1. [30] A. Vázquez-Ingelmo, J. Cruz-Benito, F. J. García-Peñalvo and M. Martín-González, "Scaffolding the OEEU's Data-Driven Ecosystem to Analyze the Employability of Spanish Graduates," in *Global Implications of Emerging Technology Trends*, F. J. García-Peñalvo, Ed. pp. 236-255, Hershey, PA: IGI Global, 2018. doi: 10.4018/978-1-5225-4944-4.ch013.
2. [416] F. García-Sánchez, J. Gómez-Isla, R. Therón, J. Cruz-Benito and J. C. Sánchez-Prieto, "Developing a Research Method to Analyze Visual Literacy Based on Cross-Cultural Characteristics," in *Global Implications of Emerging Technology Trends* pp. 19-33: IGI Global, 2018.

Books

1. [330] F. Michavila, J. M. Martínez, M. Martín-González, F. J. García-Peñalvo and J. Cruz-Benito, "Barómetro de Empleabilidad y Empleo de los Universitarios en España, 2015 (Primer informe de resultados)," Madrid, España: Observatorio de Empleabilidad y Empleo Universitarios, 2016.
2. [332] F. Michavila, J. Martínez, M. Martín-González, F. García-Peñalvo, J. Cruz-Benito and A. Vázquez-Ingelmo, "Barómetro de empleabilidad y

empleo universitarios. Edición Máster 2017," Madrid, España: Observatorio de Empleabilidad y Empleo Universitarios, 2018.

Reports

1. [323] J. Cruz-Benito, R. Therón, F. J. García-Peñalvo and M. Martín-González, "Herramienta para la validación de elementos de mejora UX/Engagement para los cuestionarios de recogida de información de egresados en el contexto del Observatorio de Empleabilidad y Empleo Universitarios (OEEU)," GRIAL Research Group. University of Salamanca, Salamanca, Spain, 2017. doi: https://doi.org/10.5281/zenodo.322575.
2. [116] J. Cruz-Benito, "Systematic literature review & mapping," GRIAL Research Group, Deparment of Computers and Automatics. University of Salamanca 2016. doi: 10.5281/zenodo.165773.

5.2.2 Software released

1. [350] J. Cruz-Benito, A. Vázquez-Ingelmo and J. C. Sánchez-Prieto, "Code repository that supports the research presented in the paper "Enabling adaptability in web forms based on user characteristics detection through A/B testing and Machine Learning"," Github, 2017. Available from: https://github.com/cbjuan/paper-ieeeAccess-2017. doi: http://doi.org/10.5281/zenodo.1009618.
2. [417] J. Cruz-Benito, "Jupyter notebook developed to support the research presented in the paper "Proposing a machine learning approach to analyze and predict employment and its factors" ", Github, 2017. Available from: https://github.com/cbjuan/paper-ijimai-ml-employability. doi: https://doi.org/10.5281/zenodo.1040464.
3. [120] J. Cruz-Benito, "Code repository that supports the research presented in the paper "Software architectures supporting HCI/HMI processes: A systematic review and mapping of the literature"," Github, 2017. Available from: https://github.com/cbjuan/slr-softwareArchitectures-HCI-HMI. doi: http://doi.org/10.5281/zenodo.1101169

5.2.3 Official records on intellectual property

1. [418] J. Cruz-Benito, C. Maderuelo Martin, F. J. García-Peñalvo, R. Therón Sánchez, A. Martín-Suárez, J. S. Pérez-Blanco, H. Zazo, J. M. Armenteros del Olmo, "Sistemas de Comunicación Bidireccional entre Mundos Virtuales y Servidores mediante Servicios Web," Spain, 2016.
2. [406] A. Martín-Suárez, C. Maderuelo Martin, J. Cruz-Benito, J. S. Pérez-Blanco, J. M. Armenteros del Olmo and H. Zazo, "Sistemas Conversacionales para el guiado de usuarios en tareas dentro de Mundo Virtuales," Spain, 2016.
3. There are another three official registers on intellectual property pending of being approved in Spain. All of them were submitted on May 2018.

5.2.4 Predoctoral research stay

1. Research center: IBM T.J. Watson Research Center
    a. Location: 200 Aqueduct Road, Ossining NY. Yorktown Heights, New York 10598 United States of America
    b. Entity: IBM Research
    c. Department: IBM Research AI & Q
    d. Supervisor: Ismael Faro Sertage
    e. Period: 20/09/2017 – 22/12/2017
    f. Topics: Human-Computer Interaction, Software Engineering, Artificial Intelligence

5.2.5 Awards

1. Best paper award in the track International Workshop on Software Engineering for E-Learning (ISELEAR'17) within the International Conference Technological Ecosystems for Enhancing Multiculturality (TEEM) 2017 held in Cádiz, Spain between October 18-20, 2017. Award granted for the paper "Improving the OEEU's data-driven technological ecosystem's interoperability with GraphQL" developed jointly to A. Vázquez-Ingelmo and F. J. García-Peñalvo.

5.2.6 Grants received

2. Grant to finance predoctoral recruitment of research staff, co-funded by the European Social Fund (EDU/310/2015)
    a. Type of grant: Predoctoral
    b. Awarding entity: Junta de Castilla y León, Spain
    c. Type: Local Government
    d. Period: 24/11/2015 – 31/12/2017
    e. Entity where activity was carried out: University of Salamanca
    f. Max. amount: 73,769.68€
3. Grant for predoctoral students enrolled in Ph.D. programs regulated by RD99/2011. Conferences modality
    a. Type of grant: Predoctoral
    b. Awarding Entity: University of Salamanca
    c. Period: 1/09/2016 – 31/08/2017
    d. Entity where activity was carried out: University of Salamanca
    e. Amount: 500€
4. Grant for predoctoral students enrolled in Ph.D. programs regulated by RD99/2011. Publications modality
    a. Type of grant: Predoctoral
    b. Awarding Entity: University of Salamanca
    c. Period: 1/09/2015 – 31/08/2016
    d. Entity where activity was carried out: University of Salamanca
    e. Amount: 500€
5. Grant for predoctoral students. Campus of International Excellence. Ph.D. School Studii Salamantinii
    a. Type of grant: Predoctoral
    b. Awarding Entity: University of Salamanca
    c. Period: 1/09/2013 – 31/08/2014
    d. Entity where activity was carried out: University of Salamanca
    e. Amount: 500€

References

[1] P. Kruchten, *The Rational Unified Process*, 3rd ed. (Object Technology Series). Boston, MA, USA: Addison-Wesley, 2004.

[2] L. Bass, P. Clements and R. Kazman, *Software Architecture in Practice*, 3rd ed. (The SEI Series in Software Engineering). Upper Saddle River, NJ, USA: Addison-Wesley, 2013.

[3] T. T. Hewett *et al.*, "ACM SIGCHI Curricula for Human-Computer Interaction," ACM, New York, NY, USA, 0897914740, 1992.

[4] A. Dix, "Human-Computer Interaction," in *Encyclopedia of Database Systems*, L. Liu and M. T. ÖZsu, Eds. pp. 1327-1331, Boston, MA: Springer US, 2009. doi: 10.1007/978-0-387-39940-9_192.

[5] J. Cannan and H. Hu, "Human-machine interaction (HMI): A survey," *University of Essex,* 2011.

[6] J.-M. Hoc, "From human–machine interaction to human–machine cooperation," *Ergonomics,* vol. 43, no. 7, pp. 833-843, 2000.

[7] C. D. Tran, H. Ezzedine and C. Kolski, "Evaluation of Agent-based Interactive Systems: Proposal of an Electronic Informer Using Petri Nets," *J. UCS,* vol. 14, no. 19, pp. 3202-3216, 2008.

[8] A. Seffah, J. Gulliksen and M. Desmarais, "An introduction to human-centered software engineering: Integrating usability in the development process," in *Human-Centered Software Engineering --- Integrating Usability in the Software Development Lifecycle*, A. Seffah, J. Gulliksen and M. Desmarais, Eds. pp. 3-14, Dordrecht: Springer Netherlands, 2005. doi: 10.1007/1-4020-4113-6_1.

[9] A. García-Holgado and F. J. García-Peñalvo, "Architectural pattern to improve the definition and implementation of eLearning ecosystems," *Science of Computer Programming,* vol. 129, no. Supplement C, pp. 20-34, 2016/11/01/ 2016. doi: https://doi.org/10.1016/j.scico.2016.03.010.

[10] A. García-Holgado and F. J. García-Peñalvo, "A Metamodel Proposal for Developing Learning Ecosystems," in *Learning and Collaboration Technologies. Novel Learning Ecosystems: 4th International Conference, LCT 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part I*, P. Zaphiris and A. Ioannou, Eds. pp. 100-109, Cham: Springer International Publishing, 2017. doi: 10.1007/978-3-319-58509-3_10.

[11] A. García-Holgado and F. García-Peñalvo, "Human interaction in learning ecosystems based on open source solutions," presented in HCI International 2018. 20th Conference on Human-Computer Interaction, Las Vegas, Nevada, USA, 15-20 July 2018, 2018.

[12] J. Bosch and N. Juristo, "Designing software architectures for usability," in *Proceedings of the 25th International Conference on Software Engineering*, 2003, pp. 757-758: IEEE Computer Society

[13] B. E. John and L. Bass, "Usability and software architecture," *Behaviour & Information Technology,* vol. 20, no. 5, pp. 329-338, 2001/01/01 2001. doi: 10.1080/01449290110081686.

[14] A. Seffah, "Adding Usability Quality Attributes into Interactive Systems Architecture: A Pattern-Based Approach," in *Patterns of HCI Design and HCI Design of Patterns: Bridging HCI Design and Model-Driven Software Engineering*, A. Seffah, Ed. pp. 59-80, Cham: Springer International Publishing, 2015. doi: 10.1007/978-3-319-15687-3_4.

[15] A. Goldberg and D. Robson, *Smalltalk-80: the language and its implementation.* Menlo Park, California, USA: Addison-Wesley Longman Publishing Co., Inc., 1983.

[16] J. Coutaz, "PAC, an object oriented model for dialog design," in *Human–Computer Interaction–INTERACT'87* pp. 431-436, Federal Republic of Germany: Elsevier, 1987.

[17] L. Bass, B. E. John and J. Kates, "Achieving usability through software architecture," presented in 23rd International Conference on Software Engineering, 2001.

[18] D. Agrawal, S. Das and A. El Abbadi, "Big data and cloud computing: current state and future opportunities," in *Proceedings of the 14th International Conference on Extending Database Technology*, Uppsala, Sweden, 2011, pp. 530-533, New York, NY, USA: ACM

[19] R. Agrawal *et al.*, "The Claremont report on database research," *ACM SIGMOD Record,* vol. 37, no. 3, pp. 9-19, 2008.
[20] M. Armbrust *et al.*, "Above the clouds: A Berkeley view of cloud computing," *Dept. Electrical Eng. and Comput. Sciences, University of California, Berkeley, Rep. UCB/EECS,* vol. 28, 2009.
[21] M. Barlow, "Innovation, Security, and Compliance in a World of Big Data," O'Reilly 2014.
[22] J. Cohen, B. Dolan, M. Dunlap, J. M. Hellerstein and C. Welton, "MAD skills: new analysis practices for big data," *Proceedings of the VLDB Endowment,* vol. 2, no. 2, pp. 1481-1492, 2009.
[23] S. Lohr, "The age of big data," *New York Times,* vol. 11, 2012.
[24] C. Lynch, "Big data: How do your data grow?," *Nature,* vol. 455, no. 7209, pp. 28-29, 2008.
[25] A. McAfee, E. Brynjolfsson, T. H. Davenport, D. J. Patil and D. Barton, "Big data. The management revolution," *Harvard Business Review,* vol. 90, no. 10, pp. 61-67, 2012.
[26] J. Seo and B. Shneiderman, "Knowledge discovery in high-dimensional data: Case studies and a user survey for the rank-by-feature framework," *Visualization and Computer Graphics, IEEE Transactions on,* vol. 12, no. 3, pp. 311-322, 2006.
[27] J. P. Shim, M. Warkentin, J. F. Courtney, D. J. Power, R. Sharda and C. Carlsson, "Past, present, and future of decision support technology," *Decision support systems,* vol. 33, no. 2, pp. 111-126, 2002.
[28] J. Srivastava, R. Cooley, M. Deshpande and P.-N. Tan, "Web usage mining: Discovery and applications of usage patterns from web data," *ACM SIGKDD Explorations Newsletter,* vol. 1, no. 2, pp. 12-23, 2000.
[29] Y. Tong, L. Chen, Y. Cheng and P. S. Yu, "Mining frequent itemsets over uncertain databases," *Proceedings of the VLDB Endowment,* vol. 5, no. 11, pp. 1650-1661, 2012.
[30] A. Vázquez-Ingelmo, J. Cruz-Benito, F. J. García-Peñalvo and M. Martín-González, "Scaffolding the OEEU's Data-Driven Ecosystem to Analyze the Employability of Spanish Graduates," in *Global Implications of Emerging Technology Trends*, F. J. García-Peñalvo, Ed. pp. 236-255, Hershey, PA: IGI Global, 2018. doi: 10.4018/978-1-5225-4944-4.ch013.
[31] H. Wickham, "Tidy data," *Journal of Statistical Software,* vol. 59, no. 10, pp. 1-23, 2014.
[32] World Wide Web Consortium (W3C). (2004, 15/03/2015). *RDF Semantics. W3C Recommendation 10 February 2004*. Available from: http://www.w3.org/TR/2004/REC-rdf-mt-20040210/.
[33] H.-c. Yang, A. Dasdan, R.-L. Hsiao and D. S. Parker, "Map-reduce-merge: simplified relational data processing on large clusters," in *Proceedings of the 2007 ACM SIGMOD international conference on Management of data*, 2007, pp. 1029-1040: ACM
[34] X. Zhu and I. Davidson, *Knowledge Discovery and Data Mining: Challenges and Realities.* Hershey, PA, USA: IGI Global, 2007.
[35] H. Chen, R. H. Chiang and V. C. Storey, "Business intelligence and analytics: from big data to big impact," *MIS quarterly,* pp. 1165-1188, 2012.
[36] F. Webster, *Theories of the information society.* New York, NY, USA: Routledge, 2014.
[37] C. Stephanidis and A. Savidis, "Universal access in the information society: methods, tools, and interaction technologies," *Universal access in the information society,* vol. 1, no. 1, pp. 40-55, 2001.
[38] UNESCO, "Toward knowledge societies," in "UNESCO World Report," Conde-sur-Noireau, France, 2005.
[39] N. Stehr, *Knowledge societies.* Canada: Wiley Online Library, 1994.
[40] P. Drucker, *Post-capitalist society.* New York, NY, USA: Routledge, 2012.
[41] F. Buschmann, *Pattern oriented software architecture: a system of patters.* West Sussex, England John Wiley & Sons 1999.
[42] J. Cruz-Benito, R. Therón and F. J. García-Peñalvo, "Software Architectures Supporting Human-Computer Interaction Analysis: A Literature Review," in *Learning and Collaboration Technologies: Third International Conference, LCT 2016, Held as Part of HCI International 2016, Toronto, ON, Canada, July 17-22, 2016, Proceedings*, P. Zaphiris and A. Ioannou, Eds. pp. 125-136, Cham: Springer International Publishing, 2016. doi: 10.1007/978-3-319-39483-1_12.
[43] J. Cruz-Benito, O. Borrás-Gené, F. J. García-Peñalvo, Á. Fidalgo Blanco and R. Therón, "Extending MOOC ecosystems using web services and software architectures," in *Proceedings*

of the XVI International Conference on Human Computer Interaction pp. 1-7, Vilanova i la Geltrú, Spain: ACM, 2015. doi: 10.1145/2829875.2829923.

[44] A. Holzinger, "Human-Computer Interaction and Knowledge Discovery (HCI-KDD): What is the benefit of bringing those two fields to work together?," in *International Conference on Availability, Reliability, and Security*, 2013, pp. 319-328: Springer

[45] J. Zyt, W. Klosgen and J. Zytkow, *Handbook of data mining and knowledge discovery.* Oxford, England: Oxford University Press, 2002.

[46] O. Maimon and L. Rokach, "Introduction to knowledge discovery and data mining," in *Data Mining and Knowledge Discovery Handbook* pp. 1-15, USA: Springer, 2009.

[47] U. Fayyad, G. Piatetsky-Shapiro and P. Smyth, "From data mining to knowledge discovery in databases," *AI magazine,* vol. 17, no. 3, p. 37, 1996.

[48] U. M. Fayyad, G. Piatetsky-Shapiro and P. Smyth, "Knowledge Discovery and Data Mining: Towards a Unifying Framework," in *KDD*, 1996, vol. 96, pp. 82-88

[49] U. M. Fayyad, "Data mining and knowledge discovery: Making sense out of data," *IEEE Expert: Intelligent Systems and Their Applications,* vol. 11, no. 5, pp. 20-25, 1996.

[50] T. Imielinski and H. Mannila, "A database perspective on knowledge discovery," *Communications of the ACM,* vol. 39, no. 11, pp. 58-64, 1996.

[51] SIGKDD: The ACM's Special Interest Group on Knowledge Discovery and Data Mining. (2018, 11/06/2018). *KDD Conferences.* Available from: http://www.kdd.org/conferences.

[52] U. Fayyad, G. Piatetsky-Shapiro and P. Smyth, "The KDD process for extracting useful knowledge from volumes of data," *Communications of the ACM,* vol. 39, no. 11, pp. 27-34, 1996.

[53] P. Domingos, *The master algorithm: How the quest for the ultimate learning machine will remake our world.* New York, NY, USA: Basic Books, 2015.

[54] G. Piatetsky-Shapiro, "Knowledge discovery in databases: 10 years after," *ACM SIGKDD Explorations Newsletter,* vol. 1, no. 2, pp. 59-61, 2000.

[55] I. Benbasat, "HCI research: Future challenges and directions," *AIS Transactions on Human-Computer Interaction,* vol. 2, no. 2, p. 1, 2010.

[56] R. Poppe, R. Rienks and B. van Dijk, "Evaluating the future of HCI: challenges for the evaluation of emerging applications," in *Artifical Intelligence for Human Computing* pp. 234-250: Springer, 2007.

[57] S. B. Shneiderman and C. Plaisant, *Designing the user interface 4 th edition.* Boston, MA, USA: Pearson Addison Wesley, 2005.

[58] B. Myers, "Challenges of HCI design and implementation," *interactions,* vol. 1, no. 1, pp. 73-83, 1994.

[59] D. A. Norman and S. W. Draper, *User Centered System Design: New perspectives on human-computer interaction.* L. Erlbaum Associates Inc, 1986.

[60] C. Jarrett and G. Gaffney, *Forms that work: Designing Web forms for usability.* Burlington, USA: Morgan Kaufmann, 2009.

[61] S. Krug, *Don't make me think!: a common sense approach to Web usability.* Berkeley, USA: Pearson Education, 2000.

[62] D. Norman, "Emotion & design: attractive things work better," *interactions,* vol. 9, no. 4, pp. 36-42, 2002.

[63] M. Seckler, S. Heinz, S. Forde, A. N. Tuch and K. Opwis, "Trust and distrust on the web: User experiences and website characteristics," *Computers in Human Behavior,* vol. 45, pp. 39-50, 2015.

[64] B. Shneiderman, "Designing trust into online experiences," *Communications of the ACM,* vol. 43, no. 12, pp. 57-59, 2000.

[65] O. W. Bertelsen, "Tertiary Artifacts at the interface," *Aesthetic computing,* pp. 357-368, 2006.

[66] A. I. Mørch, G. Stevens, M. Won, M. Klann, Y. Dittrich and V. Wulf, "Component-based technologies for end-user development," *Communications of the ACM,* vol. 47, no. 9, pp. 59-62, 2004.

[67] J. Cruz-Benito, A. Vázquez-Ingelmo, J. C. Sánchez-Prieto, R. Therón, F. J. García-Peñalvo and M. Martín-González, "Enabling Adaptability in Web Forms Based on User Characteristics Detection Through A/B Testing and Machine Learning," *IEEE Access,* vol. 6, pp. 2251-2265, 2018. doi: 10.1109/ACCESS.2017.2782678.

[68] S. Bødker, "When second wave HCI meets third wave challenges," in *Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles*, 2006, pp. 1-8: ACM

[69] C. Stephanidis *et al.*, "Toward an Information Society for All: HCI challenges and R&D recommendations," *International Journal of Human-Computer Interaction,* vol. 11, no. 1, pp. 1-28, 1999.

[70] A. Holzinger, "On knowledge discovery and interactive intelligent visualization of biomedical data," in *Proceedings of the Int. Conf. on Data Technologies and Applications DATA*, 2012, pp. 5-16

[71] A. Holzinger and I. Jurisica, "Knowledge discovery and data mining in biomedical informatics: The future is in integrative, interactive machine learning solutions," in *Interactive knowledge discovery and data mining in biomedical informatics* pp. 1-18: Springer, 2014.

[72] D. J. Patil and H. Mason, O'Reilly, Ed. *Data Driven. Creating a Data Culture.* USA: O'Reilly, 2014.

[73] M. A. Waller and S. E. Fawcett, "Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management," *Journal of Business Logistics,* vol. 34, no. 2, pp. 77-84, 2013.

[74] T. H. Davenport and D. Patil, "Data Scientist: The Sexiest Job of the 21st Century," *Harvard Business Review,* no. 10, 2012.

[75] D. J. Patil, "Data Jujitsu: The Art of Turning Data into Product," O'Reilly, USA, 2014.

[76] W. S. Cleveland, "Data science: an action plan for expanding the technical areas of the field of statistics," *International statistical review,* vol. 69, no. 1, pp. 21-26, 2001.

[77] C. A. Mattmann, "Computing: A vision for data science," *Nature,* vol. 493, no. 7433, p. 473, 2013.

[78] M. Mobjörk, "Consulting versus participatory transdisciplinarity: a refined classification of transdisciplinary research," *Futures,* vol. 42, no. 8, pp. 866-873, 2010.

[79] F. Wickson, A. L. Carew and A. W. Russell, "Transdisciplinary research: characteristics, quandaries and quality," *Futures,* vol. 38, no. 9, pp. 1046-1059, 2006.

[80] G. D. Brooks, "Cherne, L., remarks at the Discover America meeting, Brussels, June 27, 1968. as cited in "Computer science: A neglected area in schools of education" " *The Phi Delta Kappan,* vol. 53, no. 2, pp. 121-122, 1971.

[81] D. Hemmendinger, "A plea for modesty," *ACM Inroads,* vol. 1, no. 2, pp. 4-7, 2010. doi: 10.1145/1805724.1805725.

[82] A. Jaimes, N. Sebe and D. Gatica-Perez, "Human-centered computing: a multimedia perspective," in *Proceedings of the 14th ACM international conference on Multimedia*, 2006, pp. 855-864: ACM

[83] L. Bannon, "Reimagining HCI: Toward a more human-centered perspective," *interactions,* vol. 18, no. 4, pp. 50-57, 2011.

[84] M. Pantic, A. Pentland, A. Nijholt and T. S. Huang, "Human Computing and Machine Understanding of Human Behavior: A Survey," in *Artifical Intelligence for Human Computing*, Berlin, Heidelberg, 2007, pp. 47-71: Springer Berlin Heidelberg

[85] E. Folmer, M. v. Welie and J. Bosch, "Bridging patterns: An approach to bridge gaps between SE and HCI," *Information and Software Technology,* vol. 48, no. 2, pp. 69-89, 2// 2006. doi: http://dx.doi.org/10.1016/j.infsof.2005.02.005.

[86] L. Briz Ponce, "Análisis de la efectividad en las Aplicaciones m-health en dispositivos móviles dentro del ámbito de la formación médica," 2016.

[87] M. M. Catalina and G. G. Arturo, *Técnicas e instrumentos de recogida y análisis de datos.* Madrid, España: Editorial UNED, 2014.

[88] J. W. Creswell, *Research design: Qualitative, quantitative, and mixed methods approaches*, 2nd ed. California, USA: Sage Publications, Inc., 2003. doi: 10.3109/08941939.2012.723954.

[89] E. Pizarro Lucas, J. Cruz Benito and O. Gil Gonzalo, "USALSIM: Learning and Professional Practicing in a 3D Virtual World," in *2nd International Workshop on Evidence-based Technology Enhanced Learning* pp. 75-82: Springer International Publishing, 2013.

[90] J. Cruz, R. Therón, E. Pizarro and F. J. García-Peñalvo, "Análisis de Datos en Mundos Virtuales Educativos," *Actas del XV Simposio Internacional de Tecnologías de la Información y las Comunicaciones en la Educación (SINTICE13),* 2013.

[91] J. Cruz, R. Therón, E. Pizarro and F. J. García-Peñalvo, "Knowledge Discovery in Virtual Worlds Usage Data: approaching Web Mining concepts to 3D Virtual Environments," *Proceedings Fourth International Workshop on Knowledge Discovery, Knowledge Management and Decision Support (Eureka-2013),* 2013.

[92] J. Cruz-Benito, R. Therón and E. Pizarro-Lucas, "Soluciones Visuales Interactivas aplicadas a Grandes Volúmenes de Datos de Entornos 3D de Aprendizaje y Prácticas," *Avances en Informática y Automática. Séptimo Workshop,* pp. 37-54, 2013.

[93] F. J. García-Peñalvo, M. J. Rodríguez-Conde, A. M. Seoane-Pardo, M. A. Conde-González, V. Zangrando and A. García-Holgado, "GRIAL (GRupo de investigación en InterAcción y eLearning), USAL," *IE Comunicaciones: Revista Iberoamericana de Informática Educativa,* no. 15, pp. 85-94, 2012.

[94] GRIAL Research Group, "GRIAL Research Group Scientific Production Report (2011-2017). Version 2.0," Salamanca, Spain, Technical Report, GRIAL-TR-2018-004, 2018. Available from: https://goo.gl/kiUFn9. doi: 10.5281/zenodo.1217097.

[95] D. A. Gómez Aguilar, F. J. García-Peñalvo and R. Therón, "Analítica Visual en eLearning," *El Profesional de la Información,* vol. 23, no. 3, pp. 233-242, 2014.

[96] D. A. Gómez-Aguilar, Á. Hernández-García, F. J. García-Peñalvo and R. Therón, "Tap into visual analysis of customization of grouping of activities in eLearning," *Computers in Human Behavior,* vol. 47, pp. 60-67, 2015. doi: 10.1016/j.chb.2014.11.001.

[97] D. A. Gómez-Aguilar, "Analítica Visual en eLearning," Department of Computer Science and Automation, University of Salamanca, Salamanca, Spain, 2015.

[98] M. Á. Conde González, "Personalización del aprendizaje: Framework de servicios para la integración de aplicaciones online en los sistemas de gestión del aprendizaje," Department of Computer Science and Automation, University of Salamanca, Salamanca, Spain, 2012.

[99] J. García, F. J. García-Peñalvo, R. Therón and P. O. de Pablos, "Usability evaluation of a visual modelling tool for owl ontologies," *Journal of Universal Computer Science,* vol. 17, no. 9, pp. 1299-1313, 2011.

[100] A. García-Holgado and F. J. García-Peñalvo, "Architectural pattern to improve the definition and implementation of eLearning ecosystems," *Science of Computer Programming,* vol. 129, pp. 20-34, 2016.

[101] A. García-Holgado and F. J. García-Peñalvo, "Architectural pattern for the definition of eLearning ecosystems based on Open Source developments," in *Proceedings of 2014 International Symposium on Computers in Education (SIIE), Logroño, La Rioja, Spain, 12-14 Nov. 2014*, J. L. Sierra-Rodríguez, J. M. Dodero-Beardo and D. Burgos, Eds. pp. 93-98, USA: Institute of Electrical and Electronics Engineers (IEEE), 2014.

[102] A. García-Holgado and F. J. García-Peñalvo, "The evolution of the technological ecosystems: an architectural proposal to enhancing learning processes," presented in Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality, Salamanca, Spain, 2013. doi: 10.1145/2536536.2536623.

[103] J. C. Sánchez-Prieto, S. Olmos-Migueláñez and F. J. García-Peñalvo, "MLearning and pre-service teachers: An assessment of the behavioral intention using an expanded TAM model," *Computers in Human Behavior,* vol. 72, pp. 644-654, 2017.

[104] J. C. Sánchez-Prieto, S. Olmos-Migueláñez and F. J. García-Peñalvo, "Informal tools in formal contexts: Development of a model to assess the acceptance of mobile technologies among teachers," *Computers in Human Behavior,* vol. 55A, pp. 519-528, 2016. doi: 10.1016/j.chb.2015.07.002.

[105] J. C. Sánchez Prieto, S. Olmos Migueláñez and F. J. García-Peñalvo, "Understanding mobile learning: devices, pedagogical implications and research lines," *Teoría de la Educación. Educación y Cultura en la Sociedad de la Información,* vol. 15, no. 1, 2014.

[106] F. García-Sánchez, J. Gómez-Isla and R. Therón, "Formulario de evaluación por expertos de un cuestionario para la observación de la Alfabetización Visual en el público consumidor de nuevas tecnologías," GRIAL Research Group,, Zenodo, 2018. Available from: http://doi.org/10.5281/zenodo.1156294. doi: 10.5281/zenodo.1156294.

[107] F. García-Sánchez, J. Gómez-Isla and R. Therón, "Questionnaire for the observation of visual literacy in the public consumer in the new technologies," GRIAL Research Group,, Zenodo, 2018. Available from: http://doi.org/10.5281/zenodo.1210996. doi: 10.5281/zenodo.1210996.

[108] F. García-Sánchez and J. Cruz-Benito. (2018, 1/04/2018). *felicidadgsanchez/experts-evaluation-2018Questionnaire: Analysis performed on the experts assessment.* Available from: https://doi.org/10.5281/zenodo.1211000.

[109] F. García-Sánchez, J. Cruz-Benito, R. Therón and J. Gómez-Isla, "Designing and building systems and tools to analyze visual communications on social networks," presented in Third International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'15), Porto, Portugal, 2015. doi: http://dx.doi.org/10.1145/2808580.2808629.

[110] F. García-Sánchez, R. Therón and J. Gómez-Isla, "The relationships between visual communication and informal learning," presented in Third International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'15), Porto, Portugal, 2015. doi: http://dx.doi.org/10.1145/2808580.2808654.

[111] Á. F. Agudo-Peregrina, S. Iglesias-Pradas, M. Á. Conde-González and Á. Hernández-García, "Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning," *Computers in human behavior,* vol. 31, pp. 542-550, 2014.

[112] F. J. García Peñalvo, M. Á. Conde García, M. Alier Forment and M. J. Casany Guerrero, "Opening learning management systems to personal learning environments," *Journal of universal computer science: J. UCS,* vol. 17, no. 9, pp. 1222-1240, 2011.

[113] F. J. García-Peñalvo, M. Johnson, G. R. Alves, M. Minović and M. Á. Conde-González, "Informal learning recognition through a cloud ecosystem," *Future Generation Computer Systems,* vol. 32, pp. 282-294, 2014.

[114] B. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey and S. Linkman, "Systematic literature reviews in software engineering–a systematic literature review," *Information and software technology,* vol. 51, no. 1, pp. 7-15, 2009.

[115] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," Technical report, EBSE Technical Report EBSE-2007-01 2007.

[116] J. Cruz-Benito, "Systematic literature review & mapping," GRIAL Research Group, Deparment of Computers and Automatics. University of Salamanca 2016. doi: 10.5281/zenodo.165773.

[117] F. García-Peñalvo, "Mapping sistemáticos de literatura. Caso práctico de planificación usando Parsifal," Grupo GRIAL, Universidad de Salamanca, Salamanca, España, 2017. Available from: https://repositorio.grial.eu/handle/grial/1058. doi: 10.5281/zenodo.1069690

[118] B. A. Kitchenham, D. Budgen and O. Pearl Brereton, "Using mapping studies as the basis for further research – A participant-observer case study," *Information and Software Technology,* vol. 53, no. 6, pp. 638-651, 2011/06/01/ 2011. doi: https://doi.org/10.1016/j.infsof.2010.12.011.

[119] M. Petticrew and H. Roberts, *Systematic reviews in the social sciences: A practical guide.* USA: John Wiley & Sons, 2008.

[120] J. Cruz-Benito, "Code repository that supports the research presented in the paper "Software architectures supporting HCI/HMI processes: A systematic review and mapping of the literature"," Github, 2017. Available from: https://github.com/cbjuan/slr-softwareArchitectures-HCI-HMI. doi: http://doi.org/10.5281/zenodo.1101169

[121] F. W. Neiva, J. M. N. David, R. Braga and F. Campos, "Towards pragmatic interoperability to support collaboration: A systematic review and mapping of the literature," *Information and Software Technology,* vol. 72, pp. 137-150, 4// 2016. doi: http://dx.doi.org/10.1016/j.infsof.2015.12.013.

[122] A. B. Soomro, N. Salleh, E. Mendes, J. Grundy, G. Burch and A. Nordin, "The effect of software engineers' personality traits on team climate and performance: A Systematic Literature Review," *Information and Software Technology,* vol. 73, pp. 52-65, 5// 2016. doi: http://dx.doi.org/10.1016/j.infsof.2016.01.006.

[123] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman and The Prisma Group, "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement," *PLOS Medicine,* vol. 6, no. 7, p. e1000097, 2009. doi: 10.1371/journal.pmed.1000097.

[124] R. Kazman, L. Bass, G. Abowd and M. Webb, "SAAM: a method for analyzing the properties of software architectures," in *Proceedings of 16th International Conference on Software Engineering*, 1994, pp. 81-90 doi: 10.1109/ICSE.1994.296768.

[125] S. Card, A. Newell and T. P. Moran, *The psychology of human-computer interaction.* Hillsdale, NJ, USA: L. Erlbaum Associates Inc, 1983.

[126] E. L. Hutchins, J. D. Hollan and D. A. Norman, "Direct manipulation interfaces," *Human–Computer Interaction,* vol. 1, no. 4, pp. 311-338, 1985.

[127] Z. Chaczko, W. Alenazy and C. Y. Chan, "Middleware-based Software Architecture for Interactions in the Smart Learning Environment," *INNOVATION MANAGEMENT AND SUSTAINABLE ECONOMIC COMPETITIVE ADVANTAGE: FROM REGIONAL DEVELOPMENT TO GLOBAL GROWTH, VOLS I-VI, 2015,* 2015.

[128] R. Alonzo, S. Cremer, F. Mirza, S. Gowda, L. Mastromoro and D. O. Popa, "Multi-modal sensor and HMI integration with applications in personal robotics," in *SPIE Sensing Technology + Applications*, 2015, vol. 9494, p. 8: SPIE

[129] M. A. Mackin, P. T. Gonia and J. A. Lombay-Gonzalez, "An Information System prototype for analysis of astronaut/computer interaction during simulated EVA," in *Aerospace Conference, 2012 IEEE*, 2012, pp. 1-8 doi: 10.1109/AERO.2012.6187352.

[130] A. Bozzon, M. Brambilla, P. Fraternali, P. Speroni and G. Toffetti, "Applying Web-based Networking Protocols and Software Architectures for providing adaptivity, personalization, and remotization features to Industrial Human Machine Interface Applications," in *21st International Conference on Advanced Information Networking and Applications (AINA '07)*, 2007, pp. 940-947 doi: 10.1109/AINA.2007.39.

[131] G. Pires, N. Honorio, C. Lopes, U. Nunes and A. T. Almeida, "Autonomous wheelchair for disabled people," in *Industrial Electronics, 1997. ISIE '97., Proceedings of the IEEE International Symposium on*, 1997, pp. 797-801 vol.3 doi: 10.1109/ISIE.1997.648641.

[132] P. Tiefenbacher, F. Bumberger and G. Rigoll, "Evaluation of Industrial Touch Interfaces Using a Modular Software Architecture," in *Human-Computer Interaction. Theories, Methods, and Tools: 16th International Conference, HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part I*, M. Kurosu, Ed. pp. 589-600, Cham: Springer International Publishing, 2014. doi: 10.1007/978-3-319-07233-3_54.

[133] D. M. Calandra, A. Caso, F. Cutugno, A. Origlia and S. Rossi, "CoWME: a general framework to evaluate cognitive workload during multimodal interaction," presented in Proceedings of the 15th ACM on International conference on multimodal interaction, Sydney, Australia, 2013. doi: 10.1145/2522848.2522867.

[134] M. Caruso *et al.*, "My-World-in-My-Tablet: An Architecture for People with Physical Impairment," in *Human-Computer Interaction. Interaction Modalities and Techniques: 15th International Conference, HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013, Proceedings, Part IV*, M. Kurosu, Ed. pp. 637-647, Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. doi: 10.1007/978-3-642-39330-3_69.

[135] A. Bigdelou, L. Schwarz and N. Navab, "An adaptive solution for intra-operative gesture-based human-machine interaction," presented in Proceedings of the 2012 ACM international conference on Intelligent User Interfaces, Lisbon, Portugal, 2012. doi: 10.1145/2166966.2166981.

[136] F. Pittarello, "Designing a Context-Aware Architecture for Emotionally Engaging Mobile Storytelling," in *Human-Computer Interaction – INTERACT 2011: 13th IFIP TC 13 International Conference, Lisbon, Portugal, September 5-9, 2011, Proceedings, Part I*, P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque and M. Winckler, Eds. pp. 144-151, Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. doi: 10.1007/978-3-642-23774-4_14.

[137] C. Jacquet, Y. Bellik and Y. Bourda, "KUP: a model for the multimodal presentation of information in ambient intelligence," *IET Conference Proceedings,* pp. 432-439: Institution of Engineering and Technology, 2007, Available from: http://digital-library.theiet.org/content/conferences/10.1049/cp_20070403.

[138] S. Kim, J. Choi and M. Kim, "Interactive Software Architecture for Service Robots," in *2006 SICE-ICASE International Joint Conference*, 2006, pp. 4270-4275 doi: 10.1109/SICE.2006.314873.

[139] B. Parsons, P. Warner, A. White and R. Gill, "An adaptable user interface and controller for a rehabilitation robotic arm," in *Advanced Robotics, 1997. ICAR '97. Proceedings., 8th International Conference on*, 1997, pp. 919-923 doi: 10.1109/ICAR.1997.620291.

[140] L. Nigay and J. Coutaz, "A generic platform for addressing the multimodal challenge," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1995, pp. 98-105: ACM Press/Addison-Wesley Publishing Co.

[141] A. Heigemeyr and A. Harrer, "Information Management for Adaptive Automotive Human Machine Interfaces," presented in Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Seattle, WA, USA, 2014. doi: 10.1145/2667317.2667341.

[142] S. Jalaliniya, D. Mardanbegi, I. Sintos and D. G. Garcia, "EyeDroid: an open source mobile gaze tracker on Android for eyewear computers," presented in Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers, Osaka, Japan, 2015. doi: 10.1145/2800835.2804336.

[143] L. Wang and A. Canedo, "Offloading industrial human-machine interaction tasks to mobile devices and the cloud," in *Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA)*, 2014, pp. 1-4 doi: 10.1109/ETFA.2014.7005249.

[144] M. Mayer *et al.*, "Self-optimising Assembly Systems for Handling Large Components," in *Automation, Communication and Cybernetics in Science and Engineering 2011/2012*, S. Jeschke, I. Isenhardt, F. Hees and K. Henning, Eds. pp. 681-740, Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. doi: 10.1007/978-3-642-33389-7_53.

[145] T. S. de Alencar, L. R. Machado, L. de Oliveira Neris and V. P. de Almeida Neris, "Addressing the Users' Diversity in Ubiquitous Environments through a Low Cost Architecture," in *Universal Access in Human-Computer Interaction. Aging and Assistive Environments: 8th International Conference, UAHCI 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part III*, C. Stephanidis and M. Antona, Eds. pp. 439-450, Cham: Springer International Publishing, 2014. doi: 10.1007/978-3-319-07446-7_43.

[146] P. A. Rego, P. M. Moreira and L. P. Reis, "Architecture for Serious Games in Health Rehabilitation," in *New Perspectives in Information Systems and Technologies, Volume 2*, Á. Rocha, A. M. Correia, F. B. Tan and K. A. Stroetmann, Eds. pp. 307-317, Cham: Springer International Publishing, 2014. doi: 10.1007/978-3-319-05948-8_30.

[147] S. Bongartz, Y. Jin, F. Paternò, J. Rett, C. Santoro and L. D. Spano, "Adaptive User Interfaces for Smart Environments with the Support of Model-Based Languages," in *Ambient Intelligence: Third International Joint Conference, AmI 2012, Pisa, Italy, November 13-15, 2012. Proceedings*, F. Paternò, B. de Ruyter, P. Markopoulos, C. Santoro, E. van Loenen and K. Luyten, Eds. pp. 33-48, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. doi: 10.1007/978-3-642-34898-3_3.

[148] M. Vega-Barbas, I. Pau, M. Martín-Ruiz and F. Seoane, "Adaptive Software Architecture Based on Confident HCI for the Deployment of Sensitive Services in Smart Homes," *Sensors,* vol. 15, no. 4, p. 7294, 2015.

[149] B. Biel, T. Grill and V. Gruhn, "Exploring the benefits of the combination of a software architecture analysis and a usability evaluation of a mobile application," *Journal of Systems and Software,* vol. 83, no. 11, pp. 2031-2044, 2010/11/01/ 2010. doi: https://doi.org/10.1016/j.jss.2010.03.079.

[150] N. Dimakis *et al.*, "Facilitating human-centric service delivery using a pluggable service development framework," *International Journal of Ad Hoc and Ubiquitous Computing,* vol. 4, no. 3-4, pp. 223-236, 2009.

[151] A. Seffah, T. Mohamed, H. Habieb-Mammar and A. Abran, "Reconciling usability and interactive system architecture using patterns," *Journal of Systems and Software,* vol. 81, no. 11, pp. 1845-1852, 2008/11/01/ 2008. doi: https://doi.org/10.1016/j.jss.2008.04.037.

[152] J. Tessadori, M. Bisio, S. Martinoia and M. Chiappalone, "Modular Neuronal Assemblies Embodied in a Closed-Loop Environment: Toward Future Integration of Brains and Machines," (in English), *Frontiers in Neural Circuits,* Original Research vol. 6, no. 99, 2012-December-12 2012. doi: 10.3389/fncir.2012.00099.

[153] E. Folmer and J. Bosch, "Experiences with Software Architecture Analysis of Usability," *International Journal of Information Technology and Web Engineering (IJITWE),* vol. 3, no. 4, pp. 1-29, 2008. doi: 10.4018/jitwe.2008100101.

[154] I. Marsic and B. Dorohonceanu, "Flexible User Interfaces for Group Collaboration," *International Journal of Human–Computer Interaction,* vol. 15, no. 3, pp. 337-360, 2003/06/01 2003. doi: 10.1207/S15327590IJHC1503_02.

[155] W. A. König, R. Rädle and H. Reiterer, "Interactive design of multimodal user interfaces," *Journal on Multimodal User Interfaces,* journal article vol. 3, no. 3, pp. 197-213, April 01 2010. doi: 10.1007/s12193-010-0044-2.

[156] H. Olmedo, D. Escudero and V. Cardeñoso, "Multimodal interaction with virtual worlds XMMVR: eXtensible language for MultiModal interaction with virtual reality worlds," *Journal on Multimodal User Interfaces,* journal article vol. 9, no. 3, pp. 153-172, September 01 2015. doi: 10.1007/s12193-015-0176-5.

[157] B. Burger, I. Ferrané, F. Lerasle and G. Infantes, "Two-handed gesture recognition and fusion with speech to command a robot," *Autonomous Robots,* journal article vol. 32, no. 2, pp. 129-147, February 01 2012. doi: 10.1007/s10514-011-9263-y.

[158] L. Moshkina, S. Park, R. C. Arkin, J. K. Lee and H. Jung, "TAME: Time-Varying Affective Response for Humanoid Robots," *International Journal of Social Robotics,* journal article vol. 3, no. 3, pp. 207-221, August 01 2011. doi: 10.1007/s12369-011-0090-2.

[159] D. Malandrino, F. Mazzoni, D. Riboni, C. Bettini, M. Colajanni and V. Scarano, "MIMOSA: context-aware adaptation for ubiquitous web access," *Personal and Ubiquitous Computing,* journal article vol. 14, no. 4, pp. 301-320, May 01 2010. doi: 10.1007/s00779-009-0232-9.

[160] V. Srinivasan, R. R. Murphy and C. L. Bethel, "A Reference Architecture for Social Head Gaze Generation in Social Robotics," *International Journal of Social Robotics,* journal article vol. 7, no. 5, pp. 601-616, November 01 2015. doi: 10.1007/s12369-015-0315-x.

[161] M. V. D. Veloso, J. T. C. Filho and G. A. Barreto, "SOM4R: a Middleware for Robotic Applications Based on the Resource-Oriented Architecture," *Journal of Intelligent & Robotic Systems,* journal article vol. 87, no. 3, pp. 487-506, September 01 2017. doi: 10.1007/s10846-017-0504-y.

[162] A. Sutcliffe, S. Thew and P. Jarvis, "Experience with user-centred requirements engineering," *Requirements Engineering,* journal article vol. 16, no. 4, pp. 267-280, November 01 2011. doi: 10.1007/s00766-011-0118-z.

[163] C. Jacquet, Y. Bourda and Y. Bellik, "Multimodal Presentation of Information in a Mobile Context," in *Advanced Intelligent Environments*, A. D. Kameas, V. Callagan, H. Hagras, M. Weber and W. Minker, Eds. pp. 67-94, Boston, MA: Springer US, 2009. doi: 10.1007/978-0-387-76485-6_4.

[164] R. Capilla, L. Carvajal and H. Lin, "Addressing Usability Requirements in Mobile Software Development," in *Relating System Quality and Software Architecture*, R. Bahsoon, P. Eeles, R. Roshandel and M. Stal, Eds. pp. 303-324, Boston: Morgan Kaufmann, 2014. doi: https://doi.org/10.1016/B978-0-12-417009-4.00012-0.

[165] M. A. Mackin, P. T. Gonia and J. A. Lombay-González, "An Information System prototype for analysis of astronaut/computer interaction during simulated EVA," in *2012 IEEE Aerospace Conference*, 2012, pp. 1-8 doi: 10.1109/AERO.2012.6187352.

[166] L. Nigay and J. Coutaz, "A design space for multimodal systems: concurrent processing and data fusion," in *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, 1993, pp. 172-178: ACM

[167] F. Paterno, C. Santoro and L. D. Spano, "MARIA: A universal, declarative, multiple abstraction-level language for service-oriented applications in ubiquitous environments," *ACM Transactions on Computer-Human Interaction (TOCHI),* vol. 16, no. 4, p. 19, 2009.

[168] R. Schroeder, "Defining virtual worlds and virtual environments," *Journal For Virtual Worlds Research,* vol. 1, no. 1, 2008.

[169] R. Schroeder, *Possible worlds: the social dynamic of virtual reality technology.* USA: Westview Press, Inc., 1996.

[170] S. C. Baker, R. K. Wentz and M. M. Woods, "Using virtual worlds in education: Second Life® as an educational tool," *Teaching of Psychology,* vol. 36, no. 1, pp. 59-64, 2009.

[171] S. Warburton, "Second Life in higher education: Assessing the potential for and the barriers to deploying virtual worlds in learning and teaching," *British Journal of Educational Technology,* vol. 40, no. 3, pp. 414-426, 2009.

[172] N. Yee, "Motivations for play in online games," *CyberPsychology & Behavior,* vol. 9, no. 6, pp. 772-775, 2006.

[173] E. Castronova, "Virtual worlds: A first-hand account of market and society on the cyberian frontier," presented in CESifo Germany, 2001.

[174] C. A. Steinkuehler and D. Williams, "Where everybody knows your (screen) name: Online games as "third places"," *Journal of computer-mediated communication,* vol. 11, no. 4, pp. 885-909, 2006.

[175] J. Cruz-Benito *et al.*, "Usalpharma: A Software Architecture to Support Learning in Virtual Worlds," *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje,* vol. 11, no. 3, pp. 194-204, 2016. doi: 10.1109/RITA.2016.2589719.

[176] J. Cruz-Benito *et al.*, "Monitoring and feedback of Learning Processes in Virtual Worlds through analytics architectures: A real case," in *Sistemas y Tecnologías de Información. Actas de la 9ª Conferencia Ibérica de Sistemas y Tecnologías de Información*, vol. I Artículos, Á. Rocha, D. Fonseca, E. Redondo, L. P. Reis and M. P. Cota, Eds. pp. 1126-1131, Barcelona, España, June, 18-21, 2014: AISTI (Asociación Ibérica de Sistemas y Tecnologías de Información), 2014.

[177] J. Cruz-Benito, F. J. García-Peñalvo and R. Therón, "Defining generic data collectors for Learning Analytics: Facing up the heterogeneous data from heterogeneous environments," presented in International Conference on Advanced Learning Technologies (ICALT) 2014, Athens, Greece. , 2014. doi: 10.1109/ICALT.2014.108.

[178] C. Maderuelo, A. Martin-Suarez, J. S. Pérez-Blanco, H. Zazo, J. Cruz-Benito and A. Domínguez-Gil, "Facility-based inspection training in a virtual 3D laboratory," (in English), *Accreditation and Quality Assurance,* vol. 19, no. 5, pp. 403-409, 2014/10/01 2014. doi: 10.1007/s00769-014-1065-4.

[179] F. J. García-Peñalvo, J. Cruz-Benito, C. Maderuelo, J. S. Pérez-Blanco and A. Martín-Suárez, "Usalpharma: A Cloud-Based Architecture to Support Quality Assurance Training Processes in Health Area Using Virtual Worlds," *The Scientific World Journal,* vol. 2014, art. 659364, p. 10, 2014. doi: 10.1155/2014/659364.

[180] C. Beer, K. Clark and D. Jones, "Indicators of engagement," *Proceedings ascilite Sydney,* 2010.

[181] K. Krause, "Understanding and promoting student engagement in university learning communities," *Paper presented as keynote address: Engaged, Inert or Otherwise Occupied,* 2005.

[182] C.-M. Zhao and G. D. Kuh, "Adding value: Learning communities and student engagement," *Research in Higher Education,* vol. 45, no. 2, pp. 115-138, 2004.

[183] S. De Freitas and T. Neumann, "The use of 'exploratory learning'for supporting immersive learning in virtual environments," *Computers & Education,* vol. 52, no. 2, pp. 343-352, 2009.

[184] S. De Freitas, "Learning in immersive worlds," *London: Joint Information Systems Committee,* 2006.

[185] M. Virvou, G. Katsionis and K. Manos, "Combining Software Games with Education: Evaluation of its Educational Effectiveness," *Journal of Educational Technology & Society,* vol. 8, no. 2, 2005.

[186] R. Wojciechowski and W. Cellary, "Evaluation of learners' attitude toward learning in ARIES augmented reality environments," *Computers & Education,* 2013.

[187] D. Williams, "The mapping principle, and a research framework for virtual worlds," *Communication Theory,* vol. 20, no. 4, pp. 451-470, 2010.

[188] F. J. García-Peñalvo, J. Cruz-Benito and R. Therón Sánchez, "Visualización y Análisis de Datos en Mundos Virtuales Educativos: Comprendiendo la interacción de los usuarios en los entornos 3D," *ReVisión,* vol. 7, no. 2, pp. 46-59, 2014.

[189] G. Siemens and P. Long, "Penetrating the fog: Analytics in learning and education," *Educause Review,* vol. 46, no. 5, pp. 30-32, 2011.

[190] G. Siemens, "Learning analytics: envisioning a research discipline and a domain of practice," in *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 2012, pp. 4-8: ACM

[191] G. Salmon, "The future for (second) life and learning," *British Journal of Educational Technology,* vol. 40, no. 3, pp. 526-538, 2009.

[192] T. Atkinson, "Inside Linden Lab: Second Life [TM] for Educators," (in English), *TechTrends: Linking Research and Practice to Improve Learning,* vol. 52, no. 3, pp. 16-18, 2008/06/01 2008. doi: 10.1007/s11528-008-0147-8.

[193] L. Jarmon, K. Lim and B. Carpenter, "Introduction: Pedagogy, education and innovation in virtual worlds," *Journal of Virtual Worlds Research,* vol. 2, no. 1, pp. 3-4, 2009.

[194] L. Beard, D. Morra, K. Wilson and J. Keelan, "A survey of health-related activities on second life," *Journal of Medical Internet Research,* vol. 11, no. 2, 2009.

[195] M. N. K. Boulos, L. Hetherington and S. Wheeler, "Second Life: an overview of the potential of 3-D virtual worlds in medical and health education," *Health Information & Libraries Journal,* vol. 24, no. 4, pp. 233-245, 2007.

[196] M. N. K. Boulos, R. Ramloll, R. Jones and S. Toth-Cohen, "Web 3D for Public, Environmental and Occupational Health: Early Examples from Second Life®," *International journal of environmental research and public health,* vol. 5, no. 4, pp. 290-317, 2008.

[197] L. I. Kidd, S. J. Knisley and K. I. Morgan, "Effectiveness of a second life (®) simulation as a teaching strategy for undergraduate mental health nursing students," *Journal of psychosocial nursing and mental health services,* vol. 50, no. 7, pp. 28-37, 2012.

[198] A. Lee and Z. L. Berge, "Second Life in Healthcare Education: Virtual Environment's Potential to Improve Patient Safety," *Knowledge Management & E-Learning: An International Journal (KM&EL),* vol. 3, no. 1, pp. 17-23, 2011.

[199] A. Richardson, M. Hazzard, S. D. Challman, A. M. Morgenstein and J. K. Brueckner, "A “Second Life” for gross anatomy: Applications for multiuser virtual environments in teaching the anatomical sciences," *Anatomical sciences education,* vol. 4, no. 1, pp. 39-43, 2011.

[200] J. Schwaab *et al.*, "Using second life virtual simulation environment for mock oral emergency medicine examination," *Academic Emergency Medicine,* vol. 18, no. 5, pp. 559-562, 2011.

[201] J. Wiecha, R. Heyden, E. Sternthal and M. Merialdi, "Learning in a virtual world: experience with using second life for medical education," *Journal of medical Internet research,* vol. 12, no. 1, 2010.

[202] A. Martin-Suarez *et al.*, "Scientific Knowledge Transfer Training Through a Virtual World," *Journal of Information Technology Research (JITR),* vol. 7, no. 2, pp. 24-35, 2014. doi: 10.4018/jitr.2014040103.

[203] J. Cruz-Benito *et al.*, "Using software architectures to retrieve interaction information in eLearning environments," in *Computers in Education (SIIE), 2014 International Symposium on* pp. 117-120, 2014. doi: 10.1109/SIIE.2014.7017715.

[204] A. M. Suárez *et al.*, "Virtual congresses for pharmaceutical learning," presented in Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality, Salamanca, Spain, 2013. doi: 10.1145/2536536.2536549.

[205] C. Maderuelo, J. S. Pérez-Blanco, H. Zazo, J. M. Armenteros del Olmo, J. Cruz Benito and A. Martín-Suárez, "Auditing Training Practice To Postgraduate Pharmacy Students In A Virtual World," presented in 19th Annual Conference of the European Association of Faculties of Pharmacy (EAFP 2013), Ankara, Turkey, 2013.

[206] C. Maderuelo *et al.*, "Satisfacción de estudiantes de Postgrado con una práctica en Second Life," presented in Congreso Virtual Mundial de e-Learning 2012, 2012. Available from: http://www.slideshare.net/claudioclarenc/satisfaccin-de-estudiantes-de-postgrado-con-una-prctica-en-second-life.

[207] E. P. Lucas, J. Cruz–Benito and O. G. Gonzalo, "USALSIM: learning, professional practices and employability in a 3D virtual world," *International Journal of Technology Enhanced Learning,* vol. 5, no. 3, pp. 307-321, 2013.

[208] OpenSimulator. (2014, 1/02/2014). *OpenSim Web Page.* Available from: http://opensimulator.org.

[209] K. Borner and S. Penumarthy, "Social diffusion patterns in three-dimensional virtual worlds," *Information Visualization,* vol. 2, no. 3, pp. 182-198, 2003.

[210] P. R. Messinger *et al.*, "Virtual worlds — past, present, and future: New directions in social computing," *Decision Support Systems,* vol. 47, no. 3, pp. 204-228, 6// 2009. doi: http://dx.doi.org/10.1016/j.dss.2009.02.014.

[211] M. Fetscherin and C. Lattemann, "User acceptance of virtual worlds," *Journal of Electronic Commerce Research,* vol. 9, no. 3, pp. 231-242, 2008.

[212] B. Dalgarno and M. J. W. Lee, "What are the learning affordances of 3-D virtual environments?," *British Journal of Educational Technology,* vol. 41, no. 1, pp. 10-32, 2010.

[213] K. Borner and Y. C. Lin, "Visualizing chat log data collected in 3-D virtual worlds," in *Information Visualisation, 2001. Proceedings. Fifth International Conference on*, 2001, pp. 141-146: IEEE

[214] M. Schmidt and J. Laffey, "Visualizing Behavioral Data from a 3D Virtual Learning Environment: A Preliminary Study," in *System Science (HICSS), 2012 45th Hawaii International Conference on*, 2012, pp. 3387-3394: IEEE

[215] J. Cruz-Benito, R. Therón, F. J. García-Peñalvo and E. Pizarro Lucas, "Analyzing users' movements in virtual worlds: discovering engagement and use patterns," presented in Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality, Salamanca, Spain, 2013. doi: 10.1145/2536536.2536622.

[216] M. E. Bulger, R. E. Mayer, K. C. Almeroth and S. D. Blau, "Measuring learner engagement in computer-equipped college classrooms," *Journal of Educational Multimedia and Hypermedia,* vol. 17, no. 2, pp. 129-143, 2008.

[217] K. Krause and H. Coates, "Students' engagement in first-year university," *Assessment & Evaluation in Higher Education,* vol. 33, no. 5, pp. 493-505, 2008.

[218] I. Stovall, "Engagement and Online Learning," *UIS Community of Practice for E-Learning,* 2003.

[219] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS quarterly,* pp. 319-340, 1989.

[220] J. Shen and L. B. Eder, "Intentions to Use Virtual Worlds for Education," *Journal of Information Systems Education,* vol. 20, no. 2, 2009.

[221] M. Virvou and G. Katsionis, "On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE," *Computers & Education,* vol. 50, no. 1, pp. 154-178, 2008.

[222] D. A. Gómez Aguilar, R. Therón and F. J. García-Peñalvo, "Semantic Spiral Timelines Used as Support for e-Learning," *J. ucs,* vol. 15, no. 7, pp. 1526-1545, 2009.

[223] J. Shen and L. B. Eder, "Intentions to use virtual worlds for education," *Journal of Information Systems Education,* vol. 20, no. 2, p. 225, 2009.

[224] M. Varvello, S. Ferrari, E. Biersack and C. Diot, "Exploring second life," *IEEE/ACM Transactions on Networking (TON),* vol. 19, no. 1, pp. 80-91, 2011.

[225] Q. Wang, "Design and evaluation of a collaborative learning environment," *Computers & Education,* vol. 53, no. 4, pp. 1138-1146, 2009. doi: http://dx.doi.org/10.1016/j.compedu.2009.05.023.

[226] F. J. García-Peñalvo, J. Cruz-Benito, C. Maderuelo, J. S. Pérez-Blanco and A. Martín-Suárez, "Usalpharma: A Cloud-Based Architecture to Support Quality Assurance Training Processes in Health Area Using Virtual Worlds," *The Scientific World Journal,* vol. 2014, 2014.

[227] Linden Lab. (2015, 28/07/2015). *APIs and Web Services Portal.* Available from: http://wiki.secondlife.com/wiki/APIs_and_Web_Services_Portal.

[228] Linden Lab. (2015, 15/03/2015). *Second Life. Terms of Service.* Available from: http://www.lindenlab.com/tos.

[229] Á. del Blanco Aguado, J. Torrente, I. Martínez-Ortiz and B. Fernández-Manjón, "Análisis del Uso del Estándar SCORM para la Integración de Juegos Educativos," *Revista Iberoamericana de Tecnologías del/da Aprendizaje/Aprendizagem,* vol. 6, no. 3, pp. 118-127, 2011.

[230] A. del Blanco, A. Serrano, M. Freire, I. Martínez-Ortiz and B. Fernández-Manjón, "E-Learning standards and learning analytics. Can data collection be improved by using standard data models?," in *Global Engineering Education Conference (EDUCON), 2013 IEEE*, 2013, pp. 1255-1261: IEEE

[231] Django Software Foundation. (2015, 15/03/2015). *Django Web Framework.* Available from: https://www.djangoproject.com/.

[232] R. Fielding *et al.*, "Hypertext transfer protocol--HTTP/1.1," World Wide Web Consortium (W3C) 2070-1721, 1999.

[233] MongoDB. (2015, 15/03/2015). *Website.* Available from: https://www.mongodb.org/.

[234] J. Cruz-Benito *et al.*, "Using software architectures to retrieve interaction information in eLearning environments," in *Computers in Education (SIIE), 2014 International Symposium on*, 2014, pp. 117-120. doi: 10.1109/SIIE.2014.7017715.

[235] World Wide Web Consortium (W3C). (2012, 15/03/2015). *RDF Interfaces. A set of basic primitives and a low-level API for working with RDF data.* Available from: http://www.w3.org/TR/2012/NOTE-rdf-interfaces-20120705/.

[236] World Wide Web Consortium (W3C). (2014, 15/03/2015). *HTTP - Hypertext Transfer Protocol.* Available from: http://www.w3.org/Protocols/.

[237] MongoDB, "Top 5 Considerations When Evaluating NoSQL Databases," Whitepaper, 2013 2013.

[238] R. Agrawal, T. Imieliński and A. Swami, "Mining association rules between sets of items in large databases," in *ACM SIGMOD Record*, 1993, vol. 22, pp. 207-216: ACM

[239] R. Srikant, Q. Vu and R. Agrawal, "Mining Association Rules with Item Constraints," in *KDD*, 1997, vol. 97, pp. 67-73

[240] H. Mannila and H. Toivonen, "Discovering Generalized Episodes Using Minimal Occurrences," in *KDD*, 1996, vol. 96, pp. 146-151

[241] H. Mannila and H. Toivonen, "Levelwise search and borders of theories in knowledge discovery," *Data mining and knowledge discovery,* vol. 1, no. 3, pp. 241-258, 1997.

[242] F. J. García-Peñalvo, A. Fidalgo-Blanco and M. L. Sein-Echaluce, "Los MOOC: Un análisis desde una perspectiva de la innovación institucional universitaria," *La Cuestión Universitaria,* no. 9, pp. 117-135, 2017.

[243] S. Wilson, O. Liber, M. Johnson, P. Beauvoir, P. Sharples and C. Milligan, "Personal Learning Environments: Challenging the dominant design of educational systems," *Journal of E-learning and Knowledge Society,* vol. 3, no. 2, pp. 27-38, 2007.

[244] F. J. García-Peñalvo and A. M. Seoane Pardo, "Una revisión actualizada del concepto de eLearning. Décimo Aniversario," *Education in the Knowledge Society (EKS),* vol. 16, no. 1, pp. 119-144, 2015. doi: 10.14201/eks2015161119144 .

[245] B. Gros and F. J. García-Peñalvo, "Future Trends in the Design Strategies and Technological Affordances of E-Learning," in *Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy*, M. J. Spector, B. B. Lockee and M. D. Childress, Eds. pp. 1-23, Switzerland: Springer International Publishing, 2016. doi: 10.1007/978-3-319-17727-4_67-1.

[246] L. Stoicu-Tivadar, V. Stoicu-Tivadar, D. Berian, S. Drăgan, A. Serban and C. Serban, "eduCRATE-A Virtual Hospital Architecture," *Studies in health technology and informatics,* vol. 205, pp. 803-807, 2013.

[247] Á. Serrano-Laguna, J. Torrente, P. Moreno-Ger and B. Fernández-Manjón, "Application of learning analytics in educational videogames," *Entertainment Computing,* vol. 5, no. 4, pp. 313-322, 2014.

[248] Á. Serrano-Laguna *et al.*, "Learning Analytics and Educational Games: Lessons Learned from Practical Experience," in *Games and Learning Alliance* pp. 16-28, Switzerland: Springer, 2014.

[249] J. Cruz-Benito, O. Borrás-Gené, F. J. García-Peñalvo, Á. Fidalgo Blanco and R. Therón, "Detection of Non-Formal and Informal Learning in learning communities supported by social networks in the context of a Cooperative MOOC," in *Proceedings of the XVII International Symposium on Computers in Education (SIIE'15) (Setúbal, Portugal. 25-27 November 2015)*, M. d. R. r. Rodrigues, M. n. Llamas Nistal and M. Figueiredo, Eds., Portugal: IEEE, 2015.

[250] F. J. García-Peñalvo, J. Cruz-Benito, O. Borrás-Gené and Á. Fidalgo Blanco, "Evolution of the Conversation and Knowledge Acquisition in Social Networks Related to a MOOC Course," in *Learning and Collaboration Technologies*, vol. 9192, P. Zaphiris and A. Ioannou, Eds. Lecture Notes in Computer Science, pp. 470-481, Switzerland: Springer International Publishing, 2015. doi: 10.1007/978-3-319-20609-7_44.

[251] J. Cruz-Benito *et al.*, "Usalpharma: Una arquitectura software al servicio del aprendizaje en Mundos Virtuales," *VAEP RITA,* vol. 3, no. 3, pp. 148-159, 2015.

[252] D. Livingstone and J. Kemp, "Integrating web-based and 3D learning environments: Second Life meets Moodle," *Next Generation Technology-Enhanced Learning,* vol. 8, 2008.

[253] J. W. Kemp, D. Livingstone and P. R. Bloomfield, "SLOODLE: Connecting VLE tools with emergent teaching practice in Second Life," *British Journal of Educational Technology,* vol. 40, no. 3, pp. 551-555, 2009. doi: 10.1111/j.1467-8535.2009.00938.x.

[254] A. Konstantinidis, T. Tsiatsos, S. Demetriadis and A. Pomportsis, "Collaborative learning in OpenSim by utilizing sloodle," in *Telecommunications (AICT), 2010 Sixth Advanced International Conference on*, 2010, pp. 90-95: IEEE

[255] P. Kallonis and D. G. Sampson, "A 3D Virtual Classroom Simulation for Supporting School Teachers Training Based on Synectics-" Making the Strange Familiar"," in *Advanced Learning Technologies (ICALT), 2011 11th IEEE International Conference on*, 2011, pp. 4-6: IEEE

[256] J. Cruz-Benito, R. Therón, F. J. García-Peñalvo and E. Pizarro Lucas, "Discovering usage behaviors and engagement in an Educational Virtual World," *Computers in Human Behavior,* vol. 47, no. 0, pp. 18-25, 6// 2015. doi: 10.1016/j.chb.2014.11.028.

[257] L. Griebel *et al.*, "A scoping review of cloud computing in healthcare," *BMC medical informatics and decision making,* vol. 15, no. 1, p. 17, 2015.

[258] B. A. Collis, *Tele-learning in a digital world: The future of distance learning.* London, UK.: International Thomson Computer Press, 1996.

[259] F. J. García-Peñalvo, *Advances in E-Learning: Experiences and Methodologies.* Hershey, PA, USA: Information Science Reference (formerly Idea Group Reference), 2008.

[260] F. J. García-Peñalvo, "Docencia," in *Libro Blanco de la Universidad Digital 2010* J. Laviña Orueta and L. Mengual Pavón, Eds. pp. 29-61, Barcelona, España: Ariel, 2008.

[261] M. J. Rosenberg, *E-learning: Strategies for Delivering Knowledge in the Digital.* New York, NY, USA: McGraw-Hill, 2001.

[262] F. Llorens, "Campus virtuales: de gestores de contenidos a gestores de metodologías," *RED. Revista de Educación a Distancia,* no. 42, pp. 1-12, 2014.

[263] J.-M. Dodero *et al.*, "Development of E-Learning Solutions: Different Approaches, a Common Mission," *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje (IEEE RITA),* vol. 9, no. 2, pp. 72-80, 2014. doi: 10.1109/RITA.2014.2317532.

[264] F. Llorens, R. Molina, P. Compañ and R. Satorre, "Technological Ecosystem for Open Education," in *Smart Digital Futures 2014. Frontiers in Artificial Intelligence and Applications*, vol. 262, R. Neves-Silva, G. A. Tsihrintzis, V. Uskov, R. J. Howlett and L. C. Jain, Eds. pp. 706-715: IOS Press, 2014.

[265] A. J. Berlanga and F. J. García-Peñalvo, "Learning Design in Adaptive Educational Hypermedia Systems," *Journal of Universal Computer Science,* vol. 14, no. 22, pp. 3627-3647, 2008. doi: 10.3217/jucs-014-22-3627.

[266] P. Brusilovsky, "Adaptive Hypermedia," *User Modeling and User-Adapted Interaction,* vol. 11, no. 1-2, pp. 87-110, 2001.

[267] A. J. Berlanga and F. J. García-Peñalvo, "Learning Technology Specifications: Semantic Objects for Adaptive Learning Environments," *International Journal of Learning Technology,* vol. 1, no. 4, pp. 458-472, 2005. doi: 10.1504/IJLT.2005.007155.

[268] K. M. Kapp, *The gamification of learning and instruction: game-based methods and strategies for training and education.* San Francisco, CA, USA: Wiley, 2012.

[269] C. González-González and F. Blanco-Izquierdo, "Designing social videogames for educational uses," *Computers & Education,* vol. 58, no. 1, pp. 250-262, 2012. doi: 10.1016/j.compedu.2011.08.014.

[270] G. Surendeleg, U. Tudevdagva and Y. S. Kim, "The Contribution of Gamification on User Engagement in Fully Online Course," in *Creativity in Intelligent, Technologies and Data Science: First Conference, CIT&DS 2015, Volgograd, Russia, September 15–17, 2015, Proceedings*, A. Kravets, M. Shcherbakov, M. Kultsova and O. Shabalina, Eds. pp. 710-719, Cham: Springer International Publishing, 2015. doi: 10.1007/978-3-319-23766-4_56.

[271] Á. Fidalgo-Blanco, M. L. Sein-Echaluce, F. J. García-Peñalvo and M. Á. Conde, "Using Learning Analytics to improve teamwork assessment," *Computers in Human Behavior,* vol. 47, pp. 149-156, 2015. doi: 10.1016/j.chb.2014.11.050.

[272] M. Á. Conde and Á. Hernández-García, "Learning analytics for educational decision making," *Computers in Human Behavior,* vol. 47, no. 0, pp. 1-3, 6// 2015. doi: 10.1016/j.chb.2014.12.034.

[273] F. G. Martin, "Will massive open online courses change how we teach?," *Communications of the ACM,* vol. 55, no. 8, pp. 26-28, 2012. doi: 10.1145/2240236.2240246.

[274] F. J. García-Peñalvo, Á. Fidalgo Blanco and M. L. Sein-Echaluce Lacleta. (2014, 6/1/2014). *Tendencias en los MOOCs*. Available from: http://gredos.usal.es/jspui/handle/10366/125093.

[275] J. I. Aguaded Gómez, "La revolución MOOCs, ¿una nueva educación desde el paradigma tecnológico?," *Comunicar,* no. 41, pp. 7-8, 2013. doi: 10.3916/C41-2013-a1.

[276] C. Castaño Garrido, I. Maiz Olazabalaga and U. Garay Ruiz, "Diseño, motivación y rendimiento en un curso MOOC cooperativo," *Comunicar,* vol. 44, pp. 19-26, 2015. doi: 10.3916/C44-2015-02.

[277] A. Chiappe-Laverde, N. A. Hine and J. A. M. Silva, "Literatura y práctica: una revisión crítica acerca de los MOOC," *Comunicar,* no. 44, pp. 9-18, 2015. doi: 10.3916/C44-2015-01.

[278] V. J. Marsick and K. E. Watkins, "Informal and Incidental Learning," *New Directions for Adult and Continuing Education,* vol. 2001, no. 89, pp. 25-34, 2001. doi: 10.1002/ace.5.

[279] F. J. García-Peñalvo, R. Colomo-Palacios and M. D. Lytras, "Informal learning in work environments: training with the Social Web in the workplace," *Behaviour & Information Technology,* vol. 31, no. 8, pp. 753-755, 2012. doi: 10.1080/0144929X.2012.661548.

[280] H. Eshach, "Bridging in-school and out-of-school learning: Formal, non-formal, and informal education," *Journal of science education and technology,* vol. 16, no. 2, pp. 171-190, 2007.

[281] Á. Fidalgo Blanco, M. L. Sein-Echaluce Lacleta and F. J. García-Peñalvo, "Methodological Approach and technological Framework to break the current limitations of MOOC model," *Journal of Universal Computer Science,* vol. 21, no. 5, pp. 712-734, 2015.

[282] L. Lane. (2012, 10/1/2015). *Three Kinds of MOOCs.* . Available from: http://lisahistory.net/wordpress/2012/08/three-kinds-of-moocs/.

[283] D. Clark. (2013, 10/1/2015). *MOOCs: taxonomy of 8 types of MOOC.* . Available from: http://donaldclarkplanb.blogspot.com.es/2013/04/moocs-taxonomy-of-8-types-of-mooc.html.

[284] G. Conole, "Los MOOCs como tecnologías disruptivas: estrategias para mejorar la experiencia de aprendizaje y la calidad de los MOOCs," *Campus Virtuales. Revista Científica Iberoamericana de Tecnología Educativa,* vol. 2, no. 2, pp. 16-28, 2013.

[285] F. J. García-Peñalvo, D. Griffiths, M. Johnson, P. Sharples and D. Sherlock, "Problems and opportunities in the use of technology to manage informal learning," in *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality*, F. J. García-Peñalvo, Ed. pp. 573-580, New York, USA: ACM, 2014. doi: 10.1145/2669711.2669958.

[286] N. A. Christakis and J. H. Fowler, *Connected: The surprising power of our social networks and how they shape our lives.* New York, USA: Little, Brown and Company, 2009.

[287] C. Alario-Hoyos, M. Pérez-Sanagustín, C. Delgado-Kloos, H. Parada G, M. Muñoz-Organero and A. Rodríguez-de-las-Heras, "Analysing the Impact of Built-In and External Social Tools in a MOOC on Educational Technologies," in *Scaling up Learning for Sustained Impact*, vol. 8095, D. Hernández-Leo, T. Ley, R. Klamma and A. Harrer, Eds. Lecture Notes in Computer Science, pp. 5-18: Springer Berlin Heidelberg, 2013. doi: http://dx.doi.org/10.1007/978-3-642-40814-4_2.

[288] J. Sanchez, C. González, S. Alayón and P. Gonzalez, "Using social networks at university: The case of school of computer science," in *Global Engineering Education Conference (EDUCON), 2013 IEEE* pp. 492-496, USA: IEEE, 2013. doi: 10.1109/EduCon.2013.6530151.

[289] J. Mackness, S. Mak and R. Williams, "The ideals and reality of participating in a MOOC," in *7th International Conference on Networked Learning* pp. 266-275, 2010.

[290] A. McAuley, B. Stewart, G. Siemens and D. Cormier. (2010). *The MOOC model for digital practice*. Available from: http://www.elearnspace.org/Articles/MOOC_Final.pdf.

[291] K. Silius, T. Miilumaki, J. Huhtamaki, T. Tebest, J. Merilainen and S. Pohjolainen, "Students' motivations for social media enhanced studying and learning," *Knowledge Management & E-Learning: An International Journal (KM&EL),* vol. 2, no. 1, pp. 51-67, 2010.

[292] G. Siemens, "Connectivism: A learning theory for the digital age," *International journal of instructional technology and distance learning,* vol. 2, no. 1, pp. 3-10, 2005.

[293] M. Zapata-Ros, "Teorías y modelos sobre el aprendizaje en entornos conectados y ubicuos. Bases para un nuevo modelo teórico a partir de una visión crítica del “conectivismo”," *Education in the Knowledge Society (EKS),* vol. 16, no. 1, pp. 69-102, 2015. doi: 10.14201/eks201516169102.

[294] C. Evans, "Twitter for teaching: Can social media be used to enhance the process of learning?," *British Journal of Educational Technology,* vol. 45, no. 5, pp. 902-915, 2014. doi: 10.1111/bjet.12099.

[295] E. M. M. Gallardo Echenique, Luis and M. Bullen, "Students in higher education: Social and academic uses of digital technology," *RUSC. Universities and Knowledge Society Journal,* vol. 12, no. 1, pp. 25-37, 2015. doi: 10.7238/rusc.v12i1.2078.

[296] E. C. Wenger and W. M. Snyder, "Communities of practice: The organizational frontier," *Harvard business review,* vol. 78, no. 1, pp. 139-146, 2000.

[297] J. Cruz-Benito, O. Borras-Gene, F. García-Penalvo, A. Blanco and R. Theron, "Learning communities in social networks and their relationship with the MOOCs," *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje,* vol. PP, no. 99, pp. 1-1, 2017. doi: 10.1109/RITA.2017.2655218.

[298] F. J. García-Peñalvo, J. Cruz-Benito, O. Borrás-Gené and Á. Fidalgo Blanco, "Evolution of the Conversation and Knowledge Acquisition in Social Networks related to a MOOC Course," presented in HCI International, Los Angeles, CA, USA, In press, 2015.

[299] Universidad Politécnica de Madrid, Universidad de Zaragoza and Universidad de Salamanca. (2015). *iMOOC.* Available from: http://gridlab.upm.es/imooc/.

[300] M. L. Sein-Echaluce Lacleta, Á. Fidalgo-Blanco, F. J. García-Peñalvo and M. Á. Conde-González, "iMOOC Platform: Adaptive MOOCs," presented in Learning and Collaboration Technologies. Third International Conference, LCT 2016, Held as Part of HCI International 2016, Toronto, ON, Canada, July 17-22, 2016, 2016. Proceedings. doi: 10.1007/978-3-319-39483-1_35.

[301] Á. Fidalgo Blanco, F. J. García-Peñalvo and M. L. Sein-Echaluce Lacleta, "A methodology proposal for developing adaptive cMOOC," in *Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality*, F. J. García-Peñalvo, Ed. pp. 553-558, New York, USA: ACM, 2013.

[302] N. Sonwalkar, "The first adaptive MOOC: A case study on pedagogy framework and scalable cloud Architecture—Part I," *MOOCs Forum,* vol. 1, no. P, pp. 22-29, 2013. doi: 10.1089/mooc.2013.0007.

[303] J. Daniel, E. V. Cano and M. Gisbert, "The Future of MOOCs: Adaptive Learning or Business Model?," *RUSC. Universities and Knowledge Society Journal,* vol. 12, no. 1, pp. 64-73, 2015. doi: 10.7238/rusc.v12i1.2475.

[304] Á. Fidalgo Blanco, M. L. Sein-Echaluce Lacleta and F. J. García-Peñalvo, "MOOC cooperativo. Una integración entre cMOOC y xMOOC. Cooperative MOOC. An integration between cMOOC and xMOOC," in *II Congreso Internacional sobre Aprendizaje, Innovación y Competitividad, CINAIC 2013*, Á. F. Blanco and M. L. S.-E. Lacleta, Eds. no. Generic, pp. 481-486, Madrid: Fundación General de la Universidad Politécnica de Madrid, 2013.

[305] Á. Fidalgo Blanco, M. L. Sein-Echaluce Lacleta, O. Borrás Gené and F. J. García Peñalvo, "Educación en abierto: Integración de un MOOC con una asignatura académica," *Education in the Knowledge Society (EKS),* Recursos educativos abiertos; Aprendizaje en línea; Aprendizaje Social; Curso en línea abierto masivo vol. 15, no. 3, pp. 233-255, 2014.

[306] J. Mackness, M. Waite, G. Roberts and E. Lovegrove, "Learning in a small, task–oriented, connectivist MOOC: Pedagogical issues and implications for higher education," *2013,* MOOCs, Open Academic Practice, Higher Education vol. 14, no. 4, 2013-09-30 2013.

[307] F. J. García-Peñalvo, Á. Fidalgo-Blanco and M. L. Sein-Echaluce, "An adaptive hybrid MOOC model: Disrupting the MOOC concept in higher education," *Telematics and Informatics,* vol. 35, no. 4, pp. 1018-1030, 2018. doi: 10.1016/j.tele.2017.09.012.

[308] O. Borrás Gené, M. Martínez Núñez and Á. Fidalgo Blanco, "Gamification in MOOC: challenges, opportunities and proposals for advancing MOOC model," in *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality*, F. J. García-Peñalvo, Ed. pp. 215-220, New York, USA: ACM, 2014. doi: 10.1145/2669711.2669902.

[309] Á. Fidalgo-Blanco, M. L. Sein-Echaluce, F. J. García-Peñalvo and J. Esteban Escaño, "Improving the MOOC learning outcomes throughout informal learning activities," in *Proceedings of the Second International Conference on Technological Ecosystems for*

Enhancing Multiculturality, F. J. García-Peñalvo, Ed. pp. 611-617, New York, USA: ACM, 2014. doi: 10.1145/2669711.2669963.
[310] J. Huang, K. M. Thornton and E. N. Efthimiadis, "Conversational tagging in twitter," in *Proceedings of the 21st ACM conference on Hypertext and hypermedia*, 2010, pp. 173-178: ACM
[311] M. Efron, "Hashtag retrieval in a microblogging environment," in *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, 2010, pp. 787-788: ACM
[312] R. T. Fielding, "Architectural Styles and the Design of Network-based Software Architectures," Irvine, University of California, 2000.
[313] Moodle. (2014). *Web service API functions.* Available from: https://docs.moodle.org/dev/Web_service_API_functions.
[314] Requests. (2015). *Requests: HTTP for Humans.* Available from: docs.python-requests.org/.
[315] Twitter Inc. (2015). *REST APIs.* Available from: https://dev.twitter.com/rest/public.
[316] Tweepy Project. (2015). *Tweepy on Github.* Available from: https://github.com/tweepy.
[317] Google. (2015). *Google+ Platform APIs.* Available from: https://developers.google.com/+/api/.
[318] J. Cruz-Benito, O. Borrás-Gené, F. J. García-Peñalvo, Á. Fidalgo Blanco and R. Therón, "Comunidades de Aprendizaje en Redes Sociales y su Relación con los MOOC," *VAEP RITA,* vol. 4, no. 2, pp. 87-99, 2016.
[319] F. García, J. Portillo, J. Romo and M. Benito, "Nativos digitales y modelos de aprendizaje," in *SPDECE*, 2007
[320] D. A. Dillman, *Mail and internet surveys: The tailored design method.* New York, NY, USA: Wiley New York, 2000.
[321] J. Cruz-Benito, J. C. Sánchez-Prieto, A. Vázquez-Ingelmo, R. Therón, F. J. García-Peñalvo and M. Martín-González, "How different versions of layout and complexity of web forms affect users after they start it? A pilot experience," in *World Conference on Information Systems and Technologies*, 2018, pp. 971-979: Springer
[322] J. Cruz-Benito *et al.*, "Improving Success/Completion Ratio in Large Surveys: A Proposal Based on Usability and Engagement," in *Learning and Collaboration Technologies. Technology in Education: 4th International Conference, LCT 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part II*, P. Zaphiris and A. Ioannou, Eds. pp. 352-370, Cham: Springer International Publishing, 2017. doi: 10.1007/978-3-319-58515-4_28.
[323] J. Cruz-Benito, R. Therón, F. J. García-Peñalvo and M. Martín-González, "Herramienta para la validación de elementos de mejora UX/Engagement para los cuestionarios de recogida de información de egresados en el contexto del Observatorio de Empleabilidad y Empleo Universitarios (OEEU)," GRIAL Research Group. University of Salamanca, Salamanca, Spain, 2017. doi: https://doi.org/10.5281/zenodo.322575.
[324] R. Colomo-Palacios, F. J. García-Peñalvo, V. Stantchev and S. Misra, "Towards a social and context-aware mobile recommendation system for tourism," *Pervasive and Mobile Computing,* vol. 38, pp. 505-515, 2017. doi: 10.1016/j.pmcj.2016.03.001.
[325] D. A. Norman, *The design of everyday things: Revised and expanded edition.* New York, NY, USA: Basic Books, 2013.
[326] J. Bargas-Avila and O. Brenzikofer, "Simple but crucial user interfaces in the world wide web: introducing 20 guidelines for usable web form design," in *User Interfaces*, R. Mátrai, Ed. pp. 1-10, Rijeka, Croatia: Intech, 2011.
[327] S. P. Anderson, *Seductive Interaction Design: Creating Playful, Fun, and Effective User Experiences, Portable Document.* Berkeley, CA, USA: Pearson Education, 2011.
[328] G. Pruvost, T. Heinroth, Y. Bellik and W. Minker, "User interaction adaptation within ambient environments," in *Next Generation Intelligent Environments. Ambient Adaptive Systems*, S. Ultes, F. Nothdurft, T. Heinroth and W. Minker, Eds. 2nd ed. pp. 221-263, Cham: Springer, 2016. doi: 10.1007/978-3-319-23452-6_6.
[329] C. Flavián, R. Gurrea and C. Orús, "The effect of product presentation mode on the perceived content and continent quality of web sites," *Online Information Review,* vol. 33, no. 6, pp. 1103-1128, 2009.

[330] F. Michavila, J. M. Martínez, M. Martín-González, F. J. García-Peñalvo and J. Cruz-Benito, *Barómetro de Empleabilidad y Empleo de los Universitarios en España, 2015 (Primer informe de resultados)*. Madrid, España: Observatorio de Empleabilidad y Empleo Universitarios, 2016.

[331] F. Michavila, J. M. Martínez, M. Martín-González, F. J. García-Peñalvo and J. Cruz-Benito, "Empleabilidad de los titulados universitarios en España. Proyecto OEEU," *2018,* vol. 19, no. 1, pp. 21-39, 2018. doi: 10.14201/eks20181912139.

[332] F. Michavila, J. Martínez, M. Martín-González, F. García-Peñalvo, J. Cruz-Benito and A. Vázquez-Ingelmo, *Barómetro de empleabilidad y empleo universitarios. Edición Máster 2017.* Madrid, España: Observatorio de Empleabilidad y Empleo Universitarios, 2018. Available from: https://goo.gl/qK3kqo.

[333] F. Michavila, M. Martín-González, J. M. Martínez, F. J. García-Peñalvo and J. Cruz-Benito, "Analyzing the employability and employment factors of graduate students in Spain: The OEEU Information System," presented in Third International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'15), Porto, Portugal, October 7-9, 2015, 2015. doi: 10.1145/2808580.2808622.

[334] *Draft BS ENISO 9241-220 Ergonomics of human-computer interaction. Part 220: Processes for enabling, executing and assessing human-centred design within organizations*, 2016.

[335] *Ergonomics of human-system interaction — Part 125: Guidance on visual presentation of Information*, 2013.

[336] N. Bevan, J. Carter and S. Harker, "ISO 9241-11 revised: What have we learnt about usability since 1998?," in *International Conference on Human-Computer Interaction*, 2015, pp. 143-151: Springer

[337] N. Bevan and L. Spinhof, "Are guidelines and standards for web usability comprehensive?," in *International Conference on Human-Computer Interaction*, 2007, pp. 407-419: Springer

[338] S. Stieger and U.-D. Reips, "What are participants doing while filling in an online questionnaire: A paradata collection tool and an empirical study," *Computers in Human Behavior,* vol. 26, no. 6, pp. 1488-1495, 2010.

[339] Y. Xu, N. Chen, A. Fernandez, O. Sinno and A. Bhasin, "From Infrastructure to Culture: A/B Testing Challenges in Large Scale Social Networks," presented in Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australia, 2015. doi: 10.1145/2783258.2788602.

[340] E. Dixon, E. Enos and S. Brodmerkle, "A/b testing of a webpage," 2011.

[341] D. Siroker and P. Koomen, *A/B testing: The most powerful way to turn clicks into customers.* Hoboken, New Jersey. USA: John Wiley & Sons, 2013.

[342] W. McKinney, *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython.* Sebastopol, CA, USA: O'Reilly Media, Inc., 2012.

[343] W. McKinney. (2017). *Pandas, Python Data Analysis Library. 2017.* Available from: http://pandas.pydata.org/.

[344] W. McKinney, "Data structures for statistical computing in python," in *Proceedings of the 9th Python in Science Conference*, 2010, vol. 445, pp. 51-56: SciPy Austin, TX

[345] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research,* vol. 12, no. Oct, pp. 2825-2830, 2011.

[346] S. Raschka, *Python machine learning.* Birmingham, UK: Packt Publishing Ltd, 2015.

[347] T. Kluyver *et al.*, "Jupyter Notebooks-a publishing format for reproducible computational workflows," in *ELPUB*, 2016, pp. 87-90

[348] M. Ragan-Kelley *et al.*, "The Jupyter/IPython architecture: a unified view of computational research, from interactive exploration to communication and publication," in *AGU Fall Meeting Abstracts*, 2014

[349] F. Perez and B. E. Granger, "Project Jupyter: Computational narratives as the engine of collaborative data science," Technical report, Project Jupyter 2015.

[350] J. Cruz-Benito, A. Vázquez-Ingelmo and J. C. Sánchez-Prieto, "Code repository that supports the research presented in the paper "Enabling adaptability in web forms based on user characteristics detection through A/B testing and Machine Learning"," Github, 2017. Available from: https://goo.gl/Zy5WZp. doi: http://doi.org/10.5281/zenodo.1009618.

[351] L. Breiman, "Random Forests," *Machine Learning,* journal article vol. 45, no. 1, pp. 5-32, October 01 2001. doi: 10.1023/a:1010933404324.

[352] S. C. Johnson, "Hierarchical clustering schemes," *Psychometrika,* journal article vol. 32, no. 3, pp. 241-254, September 01 1967. doi: 10.1007/bf02289588.
[353] A. Zollanvari, R. C. Kizilirmak, Y. H. Kho and D. Hernández-Torrano, "Predicting Students' GPA and Developing Intervention Strategies Based on Self-Regulatory Learning Behaviors," *IEEE Access,* vol. 5, pp. 23792-23802, 2017. doi: 10.1109/ACCESS.2017.2740980.
[354] F. J. García-Peñalvo, J. Cruz-Benito, M. Martín-González, A. Vázquez-Ingelmo, J. C. Sánchez-Prieto and R. Therón, "Proposing a Machine Learning Approach to Analyze and Predict Employment and its Factors," *International Journal of Interactive Multimedia and Artificial Intelligence,* vol. In Press, 2018. doi: 10.9781/ijimai.2018.02.002.
[355] A. W. Cross, L. S. Bishop, J. A. Smolin and J. M. Gambetta, "Open quantum assembly language," *arXiv preprint arXiv:1707.03429,* 2017.
[356] A. Cross, "The IBM Q experience and QISKit open-source quantum computing software," *Bulletin of the American Physical Society,* 2018.
[357] IBM Research. (2018, 18/2/2018). *QISKit. Quantum Information Software Kit.* Available from: https://www.qiskit.org/.
[358] A. Kandala *et al.*, "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets," *Nature,* vol. 549, p. 242, 09/13/online 2017. doi: 10.1038/nature23879.
[359] D. Ristè *et al.*, "Demonstration of quantum advantage in machine learning," *npj Quantum Information,* vol. 3, no. 1, p. 16, 2017/04/13 2017. doi: 10.1038/s41534-017-0017-3.
[360] S. Papert, *New theories for new learnings.* Massachusetts Institute of Technology, Media Laboratory, Epistemology and Learning Group, 1984.
[361] S. Chandrasekhar, "Stochastic problems in physics and astronomy," *Reviews of modern physics,* vol. 15, no. 1, p. 1, 1943.
[362] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal,* vol. 27, no. 3, pp. 379-423, 1948. doi: 10.1002/j.1538-7305.1948.tb01338.x.
[363] J. Cruz-Benito, I. Faro, F. Martín-Fernández, R. Therón and F. J. García-Peñalvo, "A Deep-Learning-based proposal to aid users in Quantum Computing programming," presented in HCI International 2018, Las Vegas, USA, 2018. doi: 10.1007/978-3-319-91152-6_32.
[364] N. J. Nilsson, *Principles of artificial intelligence.* Palo Alto, CA. USA: Morgan Kaufmann, 2014.
[365] I. Goodfellow, Y. Bengio and A. Courville, *Deep learning.* Cambridge, Massachusetts, USA.: MIT press, 2016.
[366] Y. LeCun, Y. Bengio and G. Hinton, "Deep learning," *Nature,* vol. 521, no. 7553, pp. 436-444, 2015.
[367] *9241-210: 2010. Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems*, 2009.
[368] D. Castelvecchi, "Quantum cloud goes commercial," *Nature,* vol. 543, no. 7644, pp. 159-159, 2017.
[369] S. Lee and J. Choi, "Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity," *International Journal of Human-Computer Studies,* vol. 103, pp. 95-105, 2017.
[370] The Linux Information Project. (2004, 20/2/2018). *Source Code Definition.* Available from: http://www.linfo.org/source_code.html.
[371] M. Harman, "Why source code analysis and manipulation will always be important," in *Source Code Analysis and Manipulation (SCAM), 2010 10th IEEE Working Conference on*, 2010, pp. 7-19: IEEE
[372] IBM Research. (2018, 18/2/2018). *The IBM Quantum Experience.* Available from: https://www.research.ibm.com/ibm-q/.
[373] I. Sutskever, O. Vinyals and Q. V. Le, "Sequence to sequence learning with neural networks," in *Advances in neural information processing systems*, 2014, pp. 3104-3112
[374] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078,* 2014.
[375] O. Vinyals and Q. Le, "A neural conversational model," *arXiv preprint arXiv:1506.05869,* 2015.
[376] S. Robertson. (2017, 20/2/2018). *Translation with a Sequence to Sequence Network and Attention.* Available from: http://pytorch.org/tutorials/intermediate/seq2seq_translation_tutorial.html.

[377] D. Bahdanau, K. Cho and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473,* 2014.

[378] G. Neubig, "Neural machine translation and sequence-to-sequence models: A tutorial," *arXiv preprint arXiv:1703.01619,* 2017.

[379] E. Loup-Escande, R. Frenoy, G. Poplimont, I. Thouvenin, O. Gapenne and O. Megalakaki, "Contributions of mixed reality in a calligraphy learning task: Effects of supplementary visual feedback and expertise on cognitive load, user experience and gestural performance," *Computers in Human Behavior,* vol. 75, pp. 42-49, 2017.

[380] A. Paszke, S. Gross, S. Chintala and G. Chanan. (2018, 20/2/2018). *PyTorch. Tensors and Dynamic neural networks in Python with strong GPU acceleration.* Available from: https://github.com/pytorch/pytorch.

[381] IBM Research. (2018, 20/2/2018). *pytorch-seq2seq.* Available from: https://github.com/IBM/pytorch-seq2seq.

[382] S. Wiseman and A. M. Rush, "Sequence-to-sequence learning as beam-search optimization," *arXiv preprint arXiv:1606.02960,* 2016.

[383] IBM Research. (2017, 20/2/2018). *Open QASM. Gate and operation specification for quantum circuits.* . Available from: https://github.com/QISKit/openqasm.

[384] D. Deutsch and R. Jozsa, "Rapid solution of problems by quantum computation," in *Proc. R. Soc. Lond. A*, 1992, vol. 439, no. 1907, pp. 553-558: The Royal Society

[385] Open Neural Network Exchange. (2018, 20/2/2018). *ONNX Github repository.* Available from: https://github.com/onnx/onnx.

[386] J. Aurisset, M. Ramm and J. Parks. (2017, 05/05/2018). *Innovating Faster on Personalization Algorithms at Netflix Using Interleaving.* Available from: https://medium.com/netflix-techblog/interleaving-in-online-experiments-at-netflix-a04ee392ec55.

[387] A. Chandrashekar, F. Amat, J. Basilico and T. Jebara. (2017, 05/05/2018). *Artwork Personalization at Netflix.* Available from: https://medium.com/netflix-techblog/artwork-personalization-c589f074ad76.

[388] N. Govind. (2017, 05/05/2018). *A/B Testing and Beyond: Improving the Netflix Streaming Experience with Experimentation and Data Science.* Available from: https://medium.com/netflix-techblog/a-b-testing-and-beyond-improving-the-netflix-streaming-experience-with-experimentation-and-data-5b0ae9295bdf.

[389] L. Floridi and M. Taddeo, "What is data ethics?," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* 10.1098/rsta.2016.0360 vol. 374, no. 2083, 2016.

[390] L. Floridi, *The ethics of information.* Oxford University Press, 2013.

[391] C. O'Neil, *Weapons of math destruction: How big data increases inequality and threatens democracy.* Broadway Books, 2016.

[392] S. S. Chivukula, J. Brier and C. M. Gray, "Dark Intentions or Persuasion?: UX Designers' Activation of Stakeholder and User Values," in *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*, 2018, pp. 87-91: ACM

[393] M. Fansher, S. S. Chivukula and C. M. Gray, "# darkpatterns: UX Practitioner Conversations About Ethical Design," in *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, p. LBW082: ACM

[394] C. M. Gray, Y. Kou, B. Battles, J. Hoggatt and A. L. Toombs, "The Dark (Patterns) Side of UX Design," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, p. 534: ACM

[395] C. Solis and X. Wang, "A study of the characteristics of behaviour driven development," in *Software Engineering and Advanced Applications (SEAA), 2011 37th EUROMICRO Conference on*, 2011, pp. 383-387: IEEE

[396] D. North, "Introducing BDD," *Better Software,* no. March, 2006.

[397] J. Bonér, D. Farley, R. Kuhn and M. Thompson. (2014, 11/06/2018). *The reactive manifesto.* Available from: https://www.reactivemanifesto.org/.

[398] M. Román, C. Hess, R. Cerqueira, A. Ranganathan, R. H. Campbell and K. Nahrstedt, "A middleware infrastructure for active spaces," *IEEE pervasive computing,* vol. 1, no. 4, pp. 74-83, 2002.

[399] D. Maggiorini and L. Ripamonti, "Cloud Computing to Support the Evolution of Massive Multiplayer Online Games," in *ENTERprise Information Systems*, vol. 220, M. Cruz-Cunha, J. Varajão, P. Powell and R. Martinho, Eds. Communications in Computer and Information Science, pp. 101-110: Springer Berlin Heidelberg, 2011. doi: 10.1007/978-3-642-24355-4_11.

[400] B. Curtis, H. Krasner and N. Iscoe, "A field study of the software design process for large systems," *Communications of the ACM,* vol. 31, no. 11, pp. 1268-1287, 1988.

[401] C.-L. Wu and L.-C. Fu, "Design and realization of a framework for human–system interaction in smart homes," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on,* vol. 42, no. 1, pp. 15-31, 2012.

[402] M. Recabarren and M. Nussbaum, "Exploring the feasibility of web form adaptation to users' cultural dimension scores," *User Modeling and User-Adapted Interaction,* vol. 20, no. 1, pp. 87-108, 2010.

[403] S. J. Russell and P. Norvig, *Artificial intelligence: a modern approach.* Malaysia; Pearson Education Limited, 2016.

[404] A. M. Turing, "On computable numbers, with an application to the Entscheidungsproblem," *Proceedings of the London mathematical society,* vol. 2, no. 1, pp. 230-265, 1937.

[405] A. Turing, "Computing Machinery and Intelligence," *Mind,* vol. LIX, no. 236, pp. 433-460, 1950.

[406] A. Martín-Suárez, C. Maderuelo Martin, J. Cruz-Benito, J. S. Pérez-Blanco, J. M. Armenteros del Olmo and H. Zazo, "Sistemas Conversacionales para el guiado de usuarios en tareas dentro de Mundo Virtuales," Spain, 2016.

[407] J. Pearl and D. Mackenzie, *The Book of Why. The New Science of Cause and Effect* New York, NY, USA: Basic Books. Hachette Book Group, 2018.

[408] C. Stephanidis, "Adaptive techniques for universal access," *User modeling and user-adapted interaction,* vol. 11, no. 1-2, pp. 159-179, 2001.

[409] A. Vázquez-Ingelmo, J. Cruz-Benito and F. J. García-Peñalvo, "Improving the OEEU's data-driven technological ecosystem's interoperability with GraphQL," presented in Fifth International Conference Technological Ecosystems for Enhancing Multiculturality 2017 (TEEM'17), Cádiz, Spain, October 18-20, 2017, 2017.

[410] R. Therón, J. Cruz-Benito, F. García-Sánchez, R. Santamaría and F. García-Peñalvo, "Innovación en la enseñanza de la Interacción Persona-Ordenador: interfaces imaginadas, ciencia-ficción y trabajo con usuarios reales," presented in IV Congreso Internacional sobre Innovación, Aprendizaje y Competitividad, CINAIC 2017, Zaragoza, Spain, 2017.

[411] J. Gómez Isla, F. García-Sánchez, J. Cruz-Benito and C. González García, "Procesos colaborativos de crítica y reflexión para la coevaluación de proyectos artísticos de alumnos de Bellas Artes mediante el uso de las tecnologías móviles," presented in IV Congreso Internacional sobre Innovación, Aprendizaje y Competitividad, CINAIC 2017, Zaragoza, Spain, 2017.

[412] J. Cruz-Benito, O. Borrás-Gené, F. J. García-Peñalvo, Á. Fidalgo Blanco and R. Therón, "Detección de aprendizaje no formal e informal en Comunidades de Aprendizaje soportadas por Redes Sociales en el contexto de un MOOC Cooperativo," in *Actas del XVII Simposio Internacional de Informática Educativa (SIIE'15) (Setúbal, Portugal. 25-27 November 2015)*, M. d. R. r. Rodrigues, M. n. Llamas Nistal and M. Figueiredo, Eds. pp. 410-418, Portugal: Escola Superior de Educação do Instituto Politécnico de Setúbal, 2015.

[413] A. García-Holgado, J. Cruz-Benito and F. J. García-Peñalvo, "Comparative analysis of the Knowledge Management in Spanish Public Administration," presented in III Congreso Internacional sobre Aprendizaje, Innovación y Competitividad - CINAIC 2015, Madrid, 2015.

[414] A. García-Holgado, J. Cruz-Benito and F. J. García-Peñalvo, "Analysis of Knowledge Management Experiences in Spanish Public Administration," presented in Third International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'15), Porto, Portugal, 7-9, October, 2015.

[415] J. Cruz-Benito, R. Therón and F. J. García-Peñalvo, "Analytics of information flows and decision making in heterogeneous learning ecosystems," in *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality*, 2014, pp. 703-707: ACM

[416] F. García-Sánchez, J. Gómez-Isla, R. Therón, J. Cruz-Benito and J. C. Sánchez-Prieto, "Developing a Research Method to Analyze Visual Literacy Based on Cross-Cultural

Characteristics," in *Global Implications of Emerging Technology Trends* pp. 19-33: IGI Global, 2018.

[417] J. Cruz-Benito, "Jupyter notebook developed to support the research presented in the paper "Proposing a machine learning approach to analyze and predict employment and its factors" ", Github, 2017. Available from: https://github.com/cbjuan/paper-ijimai-ml-employability. doi: https://doi.org/10.5281/zenodo.1040464.

[418] J. Cruz-Benito *et al.*, "Sistemas de Comunicación Bidireccional entre Mundos Virtuales y Servidores mediante Servicios Web," Spain, 2016.

6 Appendixes

In the following appendixes are included the different papers related to the research scenarios developed during the thesis.

6.1 Appendix A. Usalpharma: A Cloud-Based Architecture to Support Quality Assurance Training Processes in Health Area Using Virtual Worlds

Hindawi Publishing Corporation
The Scientific World Journal
Volume 2014, Article ID 659364, 10 pages
http://dx.doi.org/10.1155/2014/659364

Research Article

Usalpharma: A Cloud-Based Architecture to Support Quality Assurance Training Processes in Health Area Using Virtual Worlds

Francisco J. García-Peñalvo,[1] Juan Cruz-Benito,[1] Cristina Maderuelo,[2] Jonás Samuel Pérez-Blanco,[2,3] and Ana Martín-Suárez[2,3]

[1] *GRIAL Research Group, Department of Computer Science and Automatics, University of Salamanca, Paseo de Canalejas 169, 37008 Salamanca, Spain*
[2] *Department of Pharmacy and Pharmaceutical Technology, University of Salamanca, Faculty of Pharmacy, Campus M. Unamuno, 37007 Salamanca, Spain*
[3] *IBSAL Salamanca Institute for Biomedical Research, University Hospital of Salamanca, Paseo San Vicente 58-182, 37007 Salamanca, Spain*

Correspondence should be addressed to Francisco J. García-Peñalvo; fgarcia@usal.es

Received 20 December 2013; Accepted 2 February 2014; Published 20 March 2014

Academic Editors: R. Colomo-Palacios and V. Stantchev

This paper discusses how cloud-based architectures can extend and enhance the functionality of the training environments based on virtual worlds and how, from this cloud perspective, we can provide support to analysis of training processes in the area of health, specifically in the field of training processes in quality assurance for pharmaceutical laboratories, presenting a tool for data retrieval and analysis that allows facing the knowledge discovery in the happenings inside the virtual worlds.

1. Introduction

Virtual worlds and serious games are now a resource increasingly accepted to train skills and acquire knowledge in various areas, involving both formal learning, especially in universities taking into account new technological-based pedagogical approaches [1], and informal learning, more oriented towards workplace training and personal skills development [2]. But how are they able to successfully face the interaction of thousands of users expecting to complete these tasks? This is because of their infrastructure and architecture. Virtual worlds and 3D serious games are currently impossible to maintain by the use of classical server structures, as the user interaction within them is changing and does not respond to traditional usage patterns and resource requests. Virtual worlds and online 3D games had a radical change from the acceptance and use of the concepts related to cloud computing and cloud-based architectures, where resources are allocated and released dynamically, so they give a changing service, agreeing with the type of interaction and load they have. For example, Second Life (SL) typically has design of one or more (physical or virtual) servers to support one region or island, but this allocation varies with the interaction of the various elements within the virtual ground, so that if the region has a use peak other servers with free resources can support the tasks and improve overall response against these workloads [3–8]. Definitely, virtual worlds and 3D serious games currently base their success on flexible architectures that support many tasks in a transparent and dynamic way. Building on this, this paper tries to extend these concepts to a different use in the virtual worlds, such as support of the interaction analysis, in this case focusing on supporting cases of students training processes in the field of the health sciences.

The current generation of 3D virtual worlds began to develop about ten years ago. SL was the first to achieve popularity and a high degree of development [9]. Since 2008 many other virtual worlds have emerged with which SL has been sharing its leadership. But currently, SL is still the most widely used virtual world for healthcare and higher education activities [10, 11]. All the content in SL has been generated by the users. They can even create their own objects and spaces using software provided by SL [11].

Several authors have explored SL [6, 12, 13] and found a wide range of healthcare-related activities. SL has been used for health education, community outreach, training healthcare providers, and market and promotes health services and simulations purposes. Reference sites in healthcare are Healthinfo Island, Imperial College London, Virtual Hallucinations, or Second Health London [14–17].

Virtual worlds such as SL provide unique opportunities to simulate real life scenarios and immerse the user in an environment that can be tailored to meet specific educational requirements. In these immersive learning environments, learners and teachers can interact from anywhere in the real world [18]. SL has been used to train healthcare professionals in virtual problem solving and communication, in addition to other conventional preclinical teaching methods, prior to student treating patients in the clinical setting [19, 20]. Several experiences have been developed in SL where students could safely practice communication and assessment skills with simulated patients [21], developing clinical diagnosis and decision-making skills [22, 23], training procedures [24, 25], assessment of competencies [10, 26, 27], and interacting with 3D physiological models [28, 29]. Also SL has been successfully used as a resource for health education [30] and for continuing education for practitioners [11, 31–33] and even has been used in real therapy sessions [34, 35].

On Usalpharma island from SL, property of the Department of Pharmacy and Pharmaceutical Technology from the University of Salamanca, we have been carrying out activities to develop professional skills in undergraduate and postgraduate students of pharmacy since 2010 [36]. These activities take place in various facilities such as a community pharmacy or a laboratory for drugs quality control. Pharmacists must be the health professionals responsible for providing patient care that ensures optimal medication therapy outcomes [37]. For this purpose, the future pharmacist should be formed in many different areas such as drug design, drug manufacturing, dispensing, or treatment monitoring. In the community pharmacy of our island, students can perform role playing activities to train metodologies of pharmaceutical care that involves the identification, resolution, and prevention of potential drug related problems [38, 39]. There are also spaces on the island for meetings and presentation of papers [40].

The other key facility is Usalpharma Lab. This laboratory simulates the installations, equipment, documentation, and tools like a real quality control laboratory of the pharmaceutical industry that fulfill regulations to such effects [41]. This kind of installations in the real world is very costly and unusual in universities. Usalpharma Lab has been used for training of pharmacy postgraduates in quality assurance. Highly positive results were obtained as regards both the achievement of the educational goals and student satisfaction [42]. To carry out these practices, both avatars of teachers and students meet in the virtual laboratory. The teacher guides and evaluates the student during the activity. But it would be ideal that the student could access the laboratory whenever he wanted using it as a tool for self-training. Data generated through the student activity could be used by the teacher to control and evaluate their activities, without having to be present at the same time [43]. These data could provide information of great relevance for the design of new practical activities as well as the evaluation and monitoring of the correct implementation of them.

The aim of this study is to deploy a cloud architecture that supports the needs described in a virtual world, including the mechanisms of data recovery and analysis of data for proper evaluation of the practices developed inside. This will be catalyzed in this paper by the development of a case study with a tool destined to get data and analyze the insights retrieved from the virtual world.

The paper is organized as follows: one section for proposal of cloud architecture, another section for components and workflows of the cloud system, one section about a study case where we use the cloud proposal, and finally the conclusions.

2. Cloud Architecture

As discussed in the Introduction, this paper aims to articulate a cloud architecture taking advantage of data and information generated within a virtual world like SL, so starting from what happens within the virtual world we can extract and analyze information, getting to discover knowledge that serves to gain insight of what happens or provide support in decision-making processes. This way we may tackle the isolation and lack of interoperability issues that virtual worlds in general and SL in particular present; these virtual worlds usually have a structure of cloud servers, and they generally do not allow integrating them into third party software components as shown in this paper.

For example, SL or other private virtual worlds do not provide APIs or explicit methods for interaction and exchange of information with other software platforms; they only provide in-world APIs to connect with web services or retrieve information from external websites.

For this reason, the software architecture we describe should attach or wrap the private cloud architecture from virtual world, getting by this way a unique and unambiguous process for data generation, storage and analysis.

In order to define a cloud architecture to solve problems such as those discussed above, it is necessary to consider a possible definition of the context of cloud computing.

Cloud Computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the datacenters that provide those services [44].

Taking this contextualization of the problem, when a cloud architecture is defined, it is essential to design the system so that it is flexible enough to adapt itself to any data center or server farm, regardless of location, organization, technology, or number of machines [45]. In this paper, the hardware part is not discussed, due to the fact that our cloud architecture proposal is intended to have the possibility of

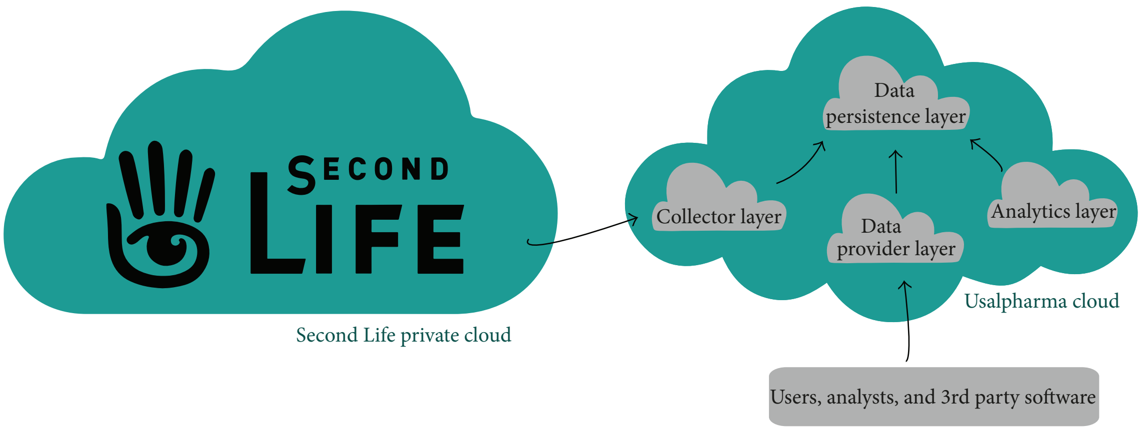


Figure 1: Usalpharma Cloud Architecture.

being deployed on any current cloud product provided by enterprises like Amazon, Google, Heroku [46–48], and so forth.

Based on the experience gained during years in the field of computing and software engineering, it is possible to base the solution design patterns and software quality factors. Our proposed architecture is based on a layer structure. The layered architecture makes each layer given different responsibilities available within the group [49], thereby forming a modular cloud system.

From these concepts we propose a modular cloud that is composed by software services (in line with SaaS concepts) [50]. Particularly, we extend the concept of software as a service (Saas), to a wider platform, obtaining concepts as CaaS (*cloud as a service*) or PaaS (*cloud platform as a service*), so each module or part of the cloud can be a cloud service or cloud architecture itself and contribute together with the other parts to achieve the common goal. As abstraction of this set of clouds within a single cloud, in the description of cloud modules, we talk about service layers or *layers*: inside these layers, we could find clouds as a service or software modules as a service.

The integration of this cloud of clouds with the Virtual World is designed thereby due to the location of SL in their own private and closed cloud of servers and resources [51, 52]. It is necessary to propose an architecture capable of interacting with it minimally, using this locked environment to generate the best final outcome for analyst, so, through minimum data output from SL private cloud, our *Usalpharma cloud* is able to maximize the final outcome and analysis capabilities. The process to maximize the outcomes with a minimal income is supported by the cloud layers, and they are responsible for the support of the whole quality assurance training process in health area. With this solution, we also try to manage the service level management in order to address dependability [53] in the privative and close approach that SL presents.

The layers for cloud architecture (named "*Usalpharma Cloud Architecture,*" Figure 1) are the following.

(i) *Data Collection Layer*. This layer of the cloud architecture is intended for retrieval and processing (prior to the storage) of the raw data which arrives from the users platform; in this case it arrives from the SL virtual world which actually is another private cloud of servers and software components.

(ii) *Data Persistence Layer*. This layer organizes and stores information persistently. It is the key layer of the architecture, on which depends the proper functioning of the cloud system. It needs to be flexible enough to accept the storage of heterogeneous data coming from the virtual world. This layer must also provide a sufficiently powerful (without bottlenecks) system so that any layer of the cloud architecture can refer simultaneously with others without affecting performance.

(iii) *Analysis Layer*. This layer is responsible for performing tasks more complex or costly in time analysis. It communicates with the persistent data layer and launches processes and algorithms to discover relationships in the data and analyze the behavior and interaction of users of virtual world.

(iv) *Data Provider Layer*. This layer is responsible for processing data requests coming from users and analysts or other 3rd party services that can be connected to our cloud. It manages interaction with the persistence layer by any data request.

3. Workflows and Components

In order to achieve a correct collaboration between software layers, or cloud modules, we need to establish which components may be the components of cloud architecture and which workflows can exist between them.

Below the components and workflows of the cloud architecture are detailed briefly, emphasizing the specific part of how the cloud environment could be built.

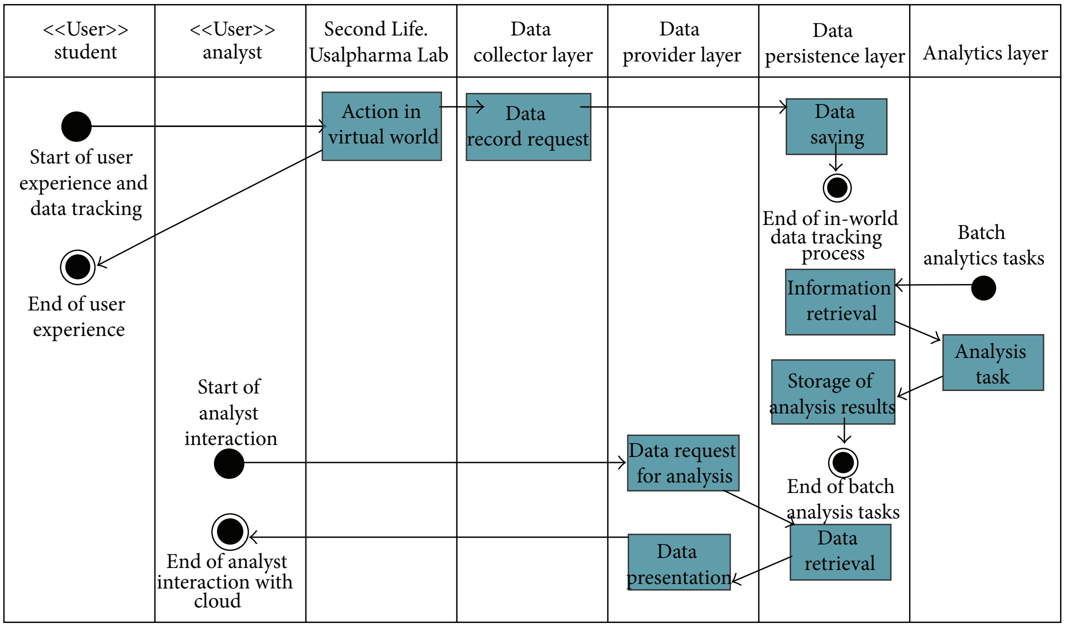


Figure 2: Activity diagram for Usalpharma cloud.

Components

(i) Web layer to collect data: the data collector layer is intended to be built based on web services, so through API requests (URI request, etc.) the data could be aggregated to cloud environment.

(ii) Database servers or cloud database to give a persistence layer and ensure access to every component in cloud architecture: this component is the central part of the cloud architecture and is used by all other components. It must guarantee the correct interaction and usage load by other components and ensure the integrity and proper preservation of the data. At present there are many database servers and storage software, each having different features, one or the other of which can be recommended depending on the technical requirements of the system. In the case of a cloud as the one proposed architecture, there are three key features that must be met; the system must ensure data persistence storage of heterogeneous data, the correct system scalability, and load balancing. Support for heterogeneous data is because the information coming from the virtual world can be of different types and vary over time depending on the events and interactions that are collected, so that the right support heterogeneous data (not only in storage but features such as indexing, search, retrieval, etc.) will ensure the full support to the training processes in the health area. The features of scalability and load balancing in data persistence layer ensure that the data growth and aggregation over time are performed correctly, thus allowing the data persistence layer is capable of supporting each once more data load and adapt to the demands of use in both the present and the future [54].

(iii) Web layer to provide information: this component should provide data on demand to the analyst or third party service. It must be able to properly access the system data persistence and above all to provide the data analyst with the way required by this want. Therefore, this component must have a specific API or software methods by which the analyst or software can define the data and the presentation that he or she needs.

(iv) Software services or cloud services to analyze data storage at database: this component is one of the most complex cloud architectures. In it lie the logic of analysis and therefore much of the usefulness of cloud architecture and the logic support of the training process that takes place in the virtual world. Within this component many algorithms and logical processes of data analysis can be found. Depending on the objectives the analysis must support different collections of methods so that they can be applied separately or together. Due to the complexity of these tasks (cost in time, CPU cycles, etc.), the executions can not be raised on demand and must be performed automatically, saving the results in the data persistence layer, so they can be picked up later by analysts or when software processes require them.

Workflows (represented in the activity diagram of Figure 2)

(a) Workflow for data collection is as follows.

(i) User interacts with the virtual world, using the 3D scene and the theoretical concepts behind this training method.

Table 1: Relation about key aspects in the training process and checkpoints that we need to know using the cloud-based analysis system.

Items to audit	Checkpoints monitored by cloud system
Laboratory access	Entry across the SAS
Installations	Entrance to all rooms
Equipment	Revision of the equipment documentation (calibration, qualification, cleaning and maintenance, etc.)
Emergency systems	Check of shower and eyebaths
	Check of emergency door, extinguisher, and the medicine chest
Documentation	SOPs (standard operating procedures)
	File cabinet of Research and Development documentation (locked)

(ii) From virtual world (SL) to data collector layer: we use a data model for interconnection between virtual world and data collector layer based on RDF specification [55]. The data model has a structure *subject + verb + predicate* (*user + action + object and interaction data* such as time). This data model allows the achievement of some goals, such as having a simple data model or having formal semantics and provable inference (e.g., for implementation of semantic analytics in the cloud) or even the possibility of using an extensible URI-based vocabulary (in the same way of part of RDF goals described in W3C specification). Requests are made via HttpRequest, containing in the URI the data as explained previously.

(iii) Data collector to persistence layer: this communication is made with the concrete API of the data persistence layer technology. The collector component composes the data with the exact method and sends it to database layer.

(b) Workflow for scheduled analytics tasks is as follows.

(i) Analytics layer executes batch actions in order to get new knowledge from data storage in databases. For this purpose, it retrieves information from database, using the software methods according to storage technology. Later it executes the analytics tasks, by applying algorithms of data mining, statistical methods, Map/Reduce strategies [56, 57], and so forth.

(ii) Finally, the results of the analysis tasks are stored in the persistence layer, so that they are recoverable by analysts.

(c) Workflow for analysts is as follows.

(i) User/software service requires data for analysis sending HttpRequests to data provider layer. This request may contain heterogeneous information about users involved in the interaction, objects, actions, date and time, and so forth. The request may also contain information about the format of the data returned from the request. This allows them to filter and get results in a more accurate way.

(ii) Data provider layer uses the data petition to compose a new petition for data persistence layer, which will return the data requested.

(iii) Finally, the data retrieved are sent to user or software in the way they require in the HttpRequest made at the beginning of the workflow.

4. Case Study

The case study presented here is based on the application of the cloud architecture described above to the case of a training process in quality assurance carried out by using a 3D environment as a virtual world like SL. Using data generated by user interaction, we propose the application of the system for collecting and analyzing data to help understand the process of training of pharmacy postgraduate students in quality assurance questions.

Our training scenario in this health area, Usalpharma Lab, is a pharmaceutical research and development (R&D) laboratory constructed in accordance with the quality standards demanded in the pharmaceutical industry. This kind of laboratories must meet current Good Laboratory Practice (GLP) norms. GLP refers to the set of regulations, operative procedures, and practices established by certain agencies, which are considered to be essential to ensure the quality and integrity of the data generated in different types of research or lines of study [58].

The layout of the laboratory is shown in Figure 3. According to this layout, the access to the laboratory is a special airlock system (SAS) that leads to the main laboratory area. Around the main area are different rooms such as document archives, storage, and production of purified and ultrapurified water. In sum it has much of the equipment and materials found in a real laboratory of this kind.

This laboratory has been built for training postgraduate students in quality assurance. Different scenarios are simulated in the installation in order to test the students auditing techniques knowledge. The students should follow a checklist with different items to audit, and they should identify regulation deviations alluding to the specific regulations aspect and should establish a classification of such deficiencies, depending on their criticality. Until now, all of the regulation deviations must be reflected in an audit report with a final decision about the adaptation level of the laboratory to the GLP regulation. Our goal is to use the cloud platform to improve the insight of the auditing process made

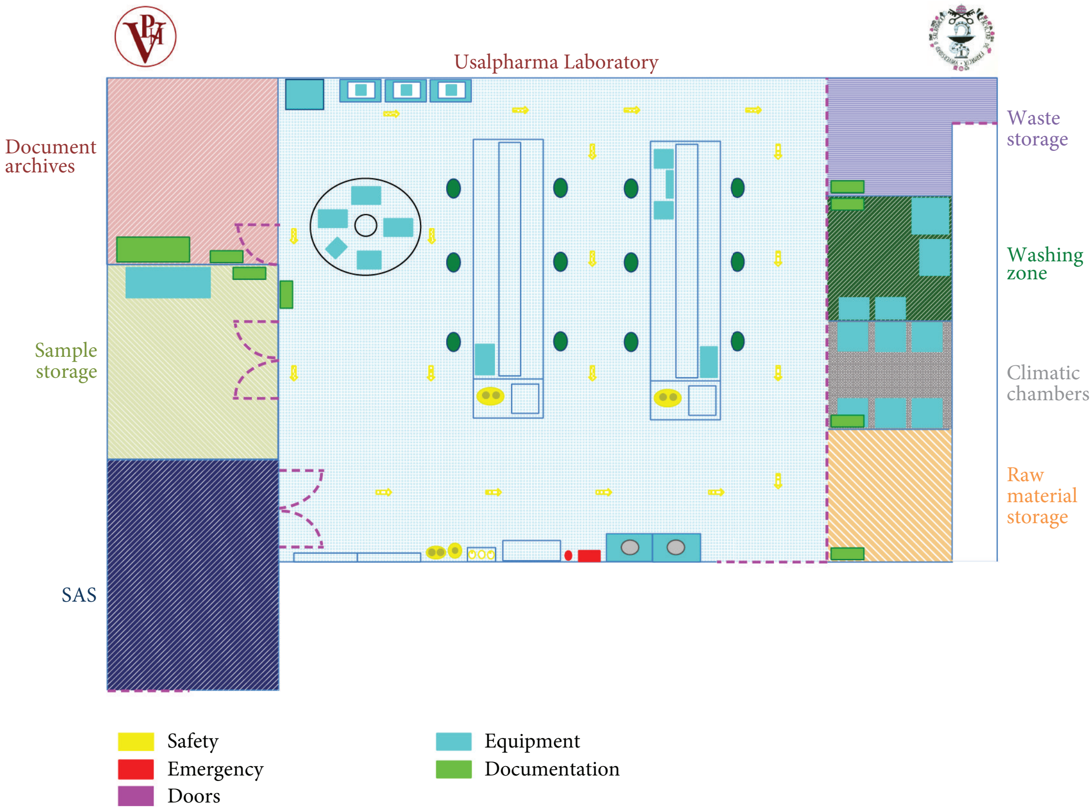


Figure 3: Usalpharma Lab layout.

Table 2: Evaluation of the training performed in Usalpharma Lab.

Evaluated aspects		Mark
From data saved in the cloud architecture	Activity control (i) Activity dates (ii) Number and duration of the lab access	Requirement to pass (i) Work in the activity period (dates fixed previously by teachers) (ii) Minimum total duration of 1 hour
	Audit methodology compliance (i) Executed actions show in a list and as a percentage (ii) Nonexecuted actions	25% of total mark (minimum required of 12.5%)
From the audit report	Ability to (i) identify regulation deviations (ii) refer the deviations to the specific regulations aspect (iii) classify the deviations depending on their criticality	60% of total mark
	Verdict about the virtual laboratory quality system	15% of total mark

by students, in order to allow these students to perform the activity without the teacher and anytime. For this reason, we propose a set of control points that will help us meet the audit system followed by them (Table 1).

Applying the cloud architecture, the teachers evaluation system could comprise both the audit report made by the student and all the interactive activity carried out for him into the platform. The report should reflect the student's assessment of the regulation compliance and the information arising by the cloud system allowing the teacher to establish the traceability of the student activity into the laboratory. So, combining the report and the cloud data, the teacher can control and evaluate the activity in a total way (Table 2).

In order to perform a proof of concept of the proposed cloud architecture, we have implemented part of it, providing a minimal system that collects data from the virtual world, performs basic analysis that allows achieving the goals proposed to obtain knowledge about interaction, and provides a system for search and retrieval of data from analysts and teachers.

To develop the proof of concept, we have made use of three main technologies: LSL (Linden Scripting Language)

Figure 4: Snapshot of a tester during the pilots.

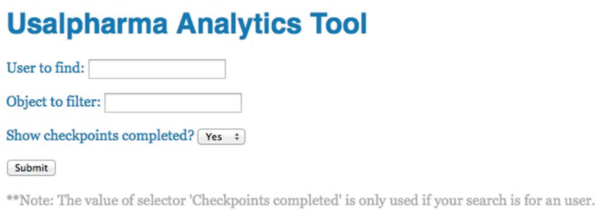


Figure 5: Web application to request data for analysis.

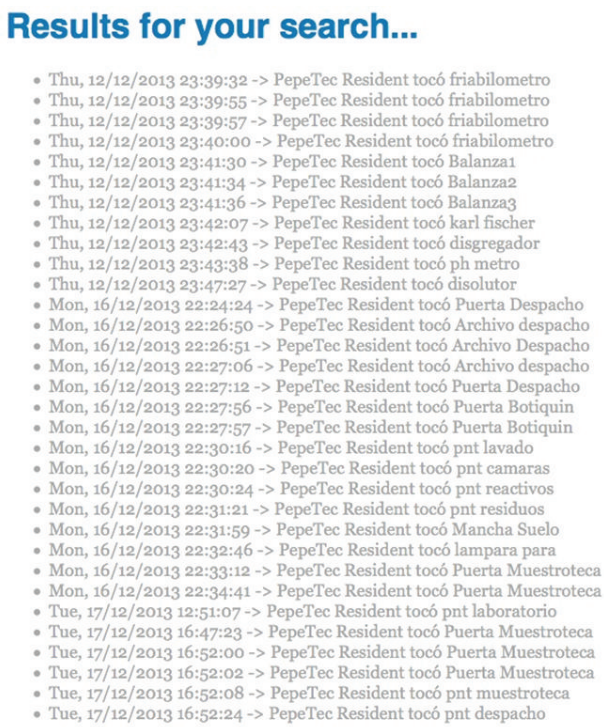


Figure 6: Results for user search.

[59], Django Web Framework [60], and MongoDB [61]. We use LSL to send data from virtual world (from the SL private cloud); this communication is made in only one way and is made by only one request; this is so, to avoid the latency in virtual world. Latency is the enemy of games or virtual worlds, and the information transmitted from the virtual world to cloud architecture must fit into a single packet wherever possible [4].

We use Django Web Framework to develop several web applications and services, which allows articulating the different components of the cloud system. Finally, we use MongoDB database to perform the data persistence layer, because this database meets the premises of support heterogeneous data, correct scalability, and load balancing [61] that we discussed in the section of *Workflows and Components*.

Once we implemented (and deployed) the proof of concept with these technologies, we performed a series of tests and pilots with users, so that they could test the proper operation of this minimal system built.

The tests were carried out for 5 postgraduate students in pharmacy, so that they perform the audit training process (Figure 4), but without having to deliver the final report on the defects or regulation deviations.

To verify proper collection and use of analysis tasks we implemented a small web tool (Figure 5) to search and retrieve results data stored in the cloud architecture (through communication with the provider data layer).

As seen in Figure 5, we could perform various types of search, filtering by user or by object, and we can select if we want to read the checkpoints achieved by the user. For example, if you run a search for a user to check their interaction with the 3D environment (without considering the evaluation of checkpoints), we can get a result as shown in Figure 6. The results include the user, the interaction performed, the object that receives the interaction, and the date and time when the action was performed.

If we select the checkpoints information in search results, we can get results quite similar to Figure 7. These results include the number and percentage of checkpoints achieved by the user and which of them (clustered by types of items to audit) the user did not complete.

In light of the results, we can say that, through this cloud architecture for recovery and analysis of data, we can get a lot of detailed information about student activity within the virtual world, which otherwise is almost impossible to obtain in virtual worlds like SL. This solves a problem observed in this type of practice in a Virtual World [43]: the teacher should be virtually present during the entire activity. Moreover, it is very difficult to review what are doing all students in the laboratory at same time. Therefore, applying this or other system to support the detailed control of user interaction, can be capital to organize activities such as this,

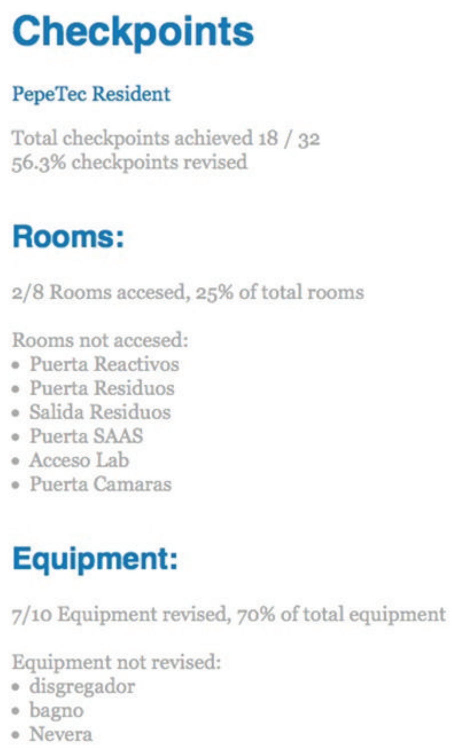


Figure 7: Search for checkpoints achieved by one user.

even increasingly complex training and allowing to obtain better results.

5. Conclusions

Throughout this paper we have specified a complete cloud architecture, starting from abstract design to workflows and components and getting to perform a proof of concept applied to a real case training audit quality assurance in the field of health care.

We consider that this cloud architecture allows carrying out the activity without the teacher's presence during the students practice. This is important to open new possibilities in training field. In addition it is possible to control and evaluate the students in a more efficient and complete way than the evaluation made in past time; this can improve the processes of training and reach more satisfactory results by the students.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

Many thanks are due to all members of Usalpharma Team, especially to Hinojal Zazo and José Manuel Armenteros del Olmo for their collaboration in the tool development. Also, special thanks are due to Felicidad García Sánchez for the fantastic design of Figure 1, *Usalpharma Cloud Architecture*. Finally, the authors would like to thank the Ministry of Science and Innovation of the Spanish Government for funding the project under Identification no. TIN2010-21695-C02-01 and the Regional Council of Education of Junta de Castilla y León (Spain) through the Projects GR47 and MPLE (ref. SA294A12-2).

References

[1] A. J. Berlanga, F. J. G. Peñalvo, and P. B. Sloep, "Towards eLearning 2.0 University," *Interactive Learning Environments*, vol. 18, no. 3, pp. 199–201, 2010.

[2] F. J. García-Peñalvo, R. Colomo-Palacios, and M. D. Lytras, "Informal learning in work environments: training with the Social Web in the workplace," *Behaviour & Information Technology*, vol. 31, no. 8, pp. 753–755, 2012.

[3] D. Maggiorini and L. A. Ripamonti, "Cloud computing to support the evolution of massive multiplayer online games," *Communications in Computer and Information Science*, vol. 220, no. 2, pp. 101–110, 2011.

[4] J. Waldo, "Scaling in games and virtual worlds," *Communications of the ACM*, vol. 51, no. 8, pp. 38–44, 2008.

[5] K. Andreas, T. Tsiatsos, T. Terzidou, and A. Pomportsis, "Fostering collaborative learning in Second Life: metaphors and affordances," *Computers and Education*, vol. 55, no. 2, pp. 603–615, 2010.

[6] M. N. K. Boulos, L. Hetherington, and S. Wheeler, "Second Life: an overview of the potential of 3-D virtual worlds in medical and health education," *Health Information and Libraries Journal*, vol. 24, no. 4, pp. 233–245, 2007.

[7] S. C. Baker, R. K. Wentz, and M. M. Woods, "Using virtual worlds in education: second life as an educational tool," *Teaching of Psychology*, vol. 36, no. 1, pp. 59–64, 2009.

[8] P. R. Messinger, E. Stroulia, K. Lyons et al., "Virtual worlds—past, present, and future: new directions in social computing," *Decision Support Systems*, vol. 47, no. 3, pp. 204–228, 2009.

[9] 2013, Second Life, http://secondlife.com.

[10] V. Patel, R. Aggarwal, D. Cohen, D. Taylor, and A. Darzi, "Implementation of an interactive virtual-world simulation for structured surgeon assessment of clinical scenarios," *Journal of the American College of Surgeons*, vol. 217, no. 2, pp. 270–279, 2013.

[11] E. Melús-Palazón, C. Bartolomé-Moreno, C. Palacín-Arbués, A. Lafuente-Lafuente et al., "Experience with using second life for medical education in a family and community medicine education unit," *BMC Medical Education*, vol. 12, no. 1, pp. 1–1, 2012.

[12] L. Beard, K. Wilson, D. Morra, and J. Keelan, "A survey of health-related activities on second life," *Journal of Medical Internet Research*, vol. 11, no. 2, article e17, 2009.

[13] K. M. N. Boulos, R. Ramloll, R. Jones, and S. Toth-Cohen, "Web 3D for public, environmental and occupational health: early examples from second life," *International journal of environmental research and public health*, vol. 5, no. 4, pp. 290–317, 2008.

[14] 2013, Healthinfo Island, http://maps.secondlife.com/secondlife/Healthinfo%20Island/128/127/23/.

[15] Imperial Collage London, http://maps.secondlife.com/secondlife/Imperial%20College%20London/149/85/26/.

[16] Virtual Hallucinations, http://maps.secondlife.com/secondlife/Sedig/26/45/21.

[17] Second Health London, http://maps.secondlife.com/secondlife/Second%20Health%20London/121/211/23/.

[18] D. R. Danforth, M. Procter, R. Chen, M. Johnson, and R. Heller, "Development of virtual patient simulations for medical education," *Journal For Virtual Worlds Research*, vol. 2, no. 2, pp. 1–1, 2009.

[19] A. Lee and Z. L. Berge, "Second life in healthcare education: virtual environment's potential to improve patient safety," *Knowledge Management & E-Learning: An International Journal*, vol. 3, no. 1, pp. 1–7, 2011.

[20] J. Phillips and Z. L. Berge, "Second life for dental education," *Journal of Dental Education*, vol. 73, no. 11, pp. 1260–1264, 2009.

[21] L. I. Kidd, S. J. Knisley, and K. I. Morgan, "Effectiveness of a second life simulation as a teaching strategy for undergraduate mental health nursing students," *Journal of Psychosocial Nursing and Mental Health Services*, vol. 50, no. 7, pp. 28–37, 2012.

[22] E. Conradi, S. Kavia, D. Burden et al., "Virtual patients in a virtual world: training paramedic students for practice," *Medical Teacher*, vol. 31, no. 8, pp. 713–720, 2009.

[23] Game-based learning for Virtual Patients—Multi patients, http://www.youtube.com/watch?v=VhQ8MjdRq4.

[24] Medical training in Second Life, http://live.wsj.com/video/medical-training-in-second-life/9F96D4FB-AFF3-4D08-8F3F-E437AF63B974.html#!9F96D4FB-AFF3-4D08-8F3F-E437AF63B974.

[25] S. Kirk, *Pharmacy Enters Second Life*, The British Journal of Clinical Pharmacy, 2009.

[26] J. Schwaab, N. Kman, R. Nagel et al., "Using second life virtual simulation environment for mock oral emergency medicine examination," *Academic Emergency Medicine*, vol. 18, no. 5, pp. 559–562, 2011.

[27] B. Zevin, R. Aggarwal, and T. P. Grantcharov, "Surgical simulation in 2013: why is it still not the standard in surgical training?" *Journal of the American College of Surgeons*, vol. 218, no. 2, pp. 294–301, 2013.

[28] A. Richardson, M. Hazzard, S. D. Challman, A. M. Morgenstein, and J. K. Brueckner, "A "Second Life" for gross anatomy: applications for multiuser virtual environments in teaching the anatomical sciences," *Anatomical Sciences Education*, vol. 4, no. 1, pp. 39–43, 2011.

[29] Genome Island, http://maps.secondlife.com/secondlife/Genome/118/144/52.

[30] V. Hall, S. Conboy-Hill, and D. Taylor, "Using virtual reality to provide health care information to people with intellectual disabilities: acceptability, usability, and potential utility," *Journal of Medical Internet Research*, vol. 13, no. 4, article e91, 2011.

[31] J. Wiecha, R. Heyden, E. Sternthal, and M. Merialdi, "Learning in a virtual world: experience with using Second Life for medical education," *Journal of Medical Internet Research*, vol. 12, no. 1, p. e1, 2010.

[32] L. Billingsley, K. Rice, M. Bennett, and S. Thibeau, "Using a multiuser virtual environment to facilitate nursing journal clubs," *Clinical Nurse Specialist*, vol. 27, no. 3, pp. 146–154, 2013.

[33] Cigna Virtual Worlds, http://www.youtube.com/watch?hl=en-GB&gl=GB&v=4KRKxv6E8pA.

[34] E. K. Yuen, J. D. Herbert, E. M. Forman, E. M. Goetter, R. Comer, and J.-C. Bradley, "Treatment of social anxiety disorder using online virtual environments in second life," *Behavior Therapy*, vol. 44, no. 1, pp. 51–61, 2013.

[35] A. Gorini, A. Gaggioli, C. Vigna, and G. Riva, "A second life for eHealth: prospects for the use of 3-D virtual worlds in clinical psychology," *Journal of Medical Internet Research*, vol. 10, no. 3, article e21, 2008.

[36] Usalpharma, http://www.youtube.com/watch?v=KhJXSzyPamc.

[37] Joint Commission of Pharmacy Practitioners, Future Vision of Pharmacy Practice, http://amcp.org/Tertiary.aspx?id=8463.

[38] A. Martín-Suárez, D. Sánchez, M. J. F. Izard, and E. Clavijo, "Aprendizaje del ejercicio profesional de la Atención Farmacéutica en un mundo virtual," *Teoría De La Educación En La Sociedad De La Información TESI*, vol. 12, no. 4, pp. 71–87, 2011.

[39] A. Martín-Suárez, E. Clavijo, D. Sánchez, M. J. García, and A. Domínguez-Gil, *Learning in A virtual World: Drug Dispensing practices in Second Life*, New Orleans, Louisiana, 2011.

[40] G. C. Millán, Z. A. Castañeda, M. J. de Jesús Valle, J. M. Lanao, F. G. López, J. S. Pérez-Blanco et al., *Utilización De Una PLataforma Virtual 3D Para La PresEntación De Trabajos En El Grado De Farmacia*, Alicante, Spain, 2013.

[41] M. C. Martín, J. S. Pérez-Blanco, H. Zazo, A. J. del Olmo, C. J. Benito, and A. Martín-Suárez, *Auditing Training Practice To Postgraduate Pharmacy Students in A Virtual World*, Anakara, Turkey, 2013.

[42] M. C. Martín, J. S. Pérez-Blanco, A. J. del Olmo, H. Zazo, D. Sánchez, C. J. Benito et al., *Satisfacción De Los Estudiantes De Postgrado Con Una Práctica De second Life*, 2012.

[43] C. J. Benito, R. Therón, E. Pizarro, and F. J. García-Peñalvo, "Knowledge discovery in virtual worlds usage data: approaching web mining concepts to 3D virtual environments," in *Proceedings 4th International Workshop on Knowledge Discovery, Knowledge Management and Decision Support*, pp. 1–11, July 2013.

[44] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski et al., "Above the clouds: a berkeley view of cloud computing. Dept electrical eng and comput sciences," Tech. Rep., University of California, Berkeley, Calif, USA, 2009.

[45] M. Armbrust, A. Fox, R. Griffith et al., "A view of cloud computing," *Communications of the ACM*, vol. 53, no. 4, pp. 50–58, 2010.

[46] Amazon Web Services (AWS), http://aws.amazon.com/.

[47] Engine Platform as a Service, https://developers.google.com/appengine/.

[48] Heroku, Cloud Application Plaftorm, https://www.heroku.com/.

[49] F. Buschmann and K. Henney, *Pattern-Oriented Software Architecture-A System of Patterns*, John Wiley and Sons, 1996.

[50] P. M. Papazoglou, "Service-oriented computing: concepts, characteristics and directions," in *Proceedings of the 4th International Conference on*, pp. 1–4, September 2003.

[51] M. Varvello, S. Ferrari, E. Biersack, and C. Diot, "Exploring second life," *IEEE/ACM Transactions on Networking*, vol. 19, no. 1, pp. 80–91, 2011.

[52] P. Mell and T. Grance, "The NIST definition of cloud computing (Draft)," *NIST Special Publication*, vol. 800, no. 145, p. 7, 2011.

[53] V. Stantchev and M. Malek, "Addressing dependability throughout the SOA life cycle," *IEEE Transactions on Services Computing*, vol. 4, no. 2, pp. 85–95, 2011.

[54] J.-D. Cryans, A. April, and A. Abran, "Criteria to compare cloud computing with current database technology," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5338, pp. 114–126, 2008.

[55] G. Klyne, J. J. Carrol, and B. McBride, *Resource Description Framework (RDF): Concepts and Abstract Syntax*, W3C Recommendation, 2004.

[56] H.-C. Yang, A. Dasdan, R.-L. Hsiao, and D. S. Parker, "Map-reduce-merge: simplified relational data processing on large clusters," in *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '07)*, pp. 1029–1040, June 2007.

[57] R. McCreadie, C. Macdonald, and I. Ounis, "MapReduce indexing strategies: studying scalability and efficiency," *Information Processing and Management*, vol. 48, no. 5, pp. 873–888, 2012.

[58] Organisation for Economic Co-operation and Development. OECD principles of good laboratory practice (as revised in 1997), Paris, OECD, 2008.

[59] Linden Scripting Language, http://wiki.secondlife.com/wiki/LSL_Portal/.

[60] Django Web Framework, https://www.djangoproject.com/.

[61] D. B. Mongo, 2013, NoSQL & Big Data Database, http://www.mongodb.com/.

6.2 Appendix B. Discovering usage behaviors and engagement in an Educational Virtual World

J. Cruz-Benito, R. Therón, F. J. García-Peñalvo and E. Pizarro Lucas, "Discovering usage behaviors and engagement in an Educational Virtual World," Computers in Human Behavior, vol. 47, no. 0, pp. 18-25, 6// 2015. doi: 10.1016/j.chb.2014.11.028.

Discovering usage behaviors and engagement in an Educational Virtual World

Abstract

This paper explores data retrieved from an Educational Virtual World to identify and validate behavior and usage patterns and engagement indicators. This data exploration is intended not to validate pre-defined questions or specific goals (for example, a comparison between engagement and academic scores) but to discover usage trends and obtain insights about users' usage of the system and their knowledge of and proficiency with the available resources and features. The engagement indicators and knowledge obtained from the analysis of these indicators regarding the users of the system, their desires, and their competencies with virtual resources will facilitate decision-making and planning by managers of the Virtual World to improve system adoption and learning effectiveness, correct usage mistakes, perform actions to enhance user exploitation of available features, and provide information to users on system usage. This knowledge and the actions based on it are capital in an eLearning ecosystem such as an Educational Virtual World, where students are able to perform tasks in 3D at any time or location without supervision.

Highlights

- Usage behaviors and engagement can be determined from user interactions.
- User engagement in Virtual Worlds may depend on the system or content.
- Users can perceive/use these systems as educational tools or for social interaction.
- Usage patterns can be identified to provide managers with deep knowledge regarding users.
- User behaviors and engagement can be used to improve learning in Virtual Worlds.

Keywords

usage behavior, usage patterns, behavior patterns, engagement indicators, Educational Virtual Worlds

1. Introduction

User perception and engagement with technology is a key aspect of a successful technological system, regardless of application. Information about user perception and engagement can be used by system developers, administrators, designers or managers to adapt, enhance or modify the technology to improve user acceptability, interest, feedback or performance. This type of research and analysis has a major presence in the field of

eLearning. Within this field of online learning, behavioral and engagement analysis could be used to propose, implement and correct various types of platforms that enable education and skills acquisition via the use of Internet-based technologies. Many authors (Beer, Clark, & Jones, 2010; Krause, 2005; Zhao & Kuh, 2004) have explored the field of behavior analysis and engagement related to online learning environments as a tool to determine how users *feel* and use technology to obtain insights into the acceptance and adaptation of these eLearning ecosystems in the Learning Process and increase success.
Tracing behavior patterns and measuring engagement on these platforms enables the determination of user interest in certain features or content and whether these features are exploited properly. These measurements enable platform managers to make decisions, correct unexpected uses, promote specific content, perform actions to avoid dropouts, improve system adoption, and adapt the structures or content of eLearning platforms to users.

An educational environment appropriate for this type of study is an Educational Virtual World. This type of environment provides an interesting field of study of user behavior and engagement because it is based on user interaction with a 3D environment and with other users. This feature, combined with the representation of the user via a virtual *alter ego* and features such as text-based chats, voice-based chats or movement between different islands or lands, allows users to express themselves very differently compared to other educational environments (De Freitas, 2006; De Freitas & Neumann, 2009). Thus, user preferences and desires in the field of Learning are often reflected more accurately than in other eLearning ecosystems.

The goals of this research were to measure engagement indicators specific to the Educational Virtual World, identify user behavior patterns (e.g., the correct use of resources and dropout patterns), determine the relationship between engagement factors and the behavior patterns identified in the proposed case study, and relate the observed rules and patterns to possible actions and decisions by Educational Virtual World managers (in case of usage error patterns, dropouts, etc.).

To achieve these goals, this research aims to conduct an exploratory analysis of a dataset retrieved from an Educational Virtual World to identify usage patterns, engagement factors and user behaviors to provide insights into the perceptions and motivations of users about a technology such as an Educational Virtual World.

The paper is divided into the following sections. The first section (Introduction) introduces the problem and concepts that will be discussed in the paper. The second section (Materials and Methods) presents the data that will be analyzed and describes the retrieval process. The third section (Theory and Calculation) provides deeper insight into the variables used in the analysis and the challenges that this analysis attempts to address. The fourth section (Results) describes the behaviors and patterns detected in the exploratory data analysis, while the fifth section (Discussion) discusses these behaviors and patterns, validating or rejecting the results as appropriate. Finally, the sixth section (Conclusions) presents several conclusions regarding the research and potential future extensions.

2. Materials and Methods

2.1 Facilities

To develop this research work, we used an Educational Virtual World developed by the University of Salamanca. This Virtual World, called USALSIM (Lucas, Cruz–Benito, & Gonzalo, 2013), is designed to provide virtual practices and *immersive* learning experiences through a 3D environment. In 2012, a series of 3D environments permitting the development of professional practices and learning in the following areas of knowledge were incorporated into USALSIM: Pharmacy, Biology and Biotechnology, Law, Humanities and Chemistry. The type of activity varies depending on the specific needs of each area of knowledge and practice; some require personal work by each user (pure interaction with 3D scenarios inside the Virtual World), while others require more discussion or role-playing tasks (interaction with other users).

This virtual world was developed under the platform known as OpenSimulator or OpenSim (OpenSimulator, 2014). This open-source platform allows the managers of the Virtual World (in this case, developers at the University of Salamanca) to control all data related to users and the 3D environment, enabling a thorough analysis of the selected data, in contrast to other Virtual Worlds (like Second Life for example) that do not allow the retrieval and analysis of detail in such detail.

Microsoft Excel was used to extract raw data (only organized in parameters and not modified) from the Virtual World for exploratory analysis. Using this software, we performed a manual analysis based on clustering of common features and measures of users and the use of basic statistics to analyze the measures. Automatic methods were not used for the analysis because this study was primarily an exploratory analysis and was intended not to address pre-defined questions but to develop findings related to the research goals.

2.2 Data Analyzed

For this study, data representing different characteristics and areas of usage were retrieved from the USALSIM Educational Virtual World and analyzed separately and in combination to identify user interests, behavior patterns or engagement measures.

As many authors have noted, Virtual Worlds feature a strong social use component (Borner & Penumarthy, 2003; Messinger et al., 2009), and even in those Virtual Worlds that implement learning activities, users often interact with each other, whether required by the educational task (such as role-playing activities) or not.

The data used were collected in two months, between November and December 2012, during testing prior to deploying this Virtual World to the general student population at the University of Salamanca. These preliminary tests were intended to study student acceptance of the Virtual World (Fetscherin & Lattemann, 2008), its relevance to education in the University and its effectiveness for skills acquisition and content learning. After the tests, both teachers and students assessed aspects of appropriateness and acceptability through surveys (Lucas et al., 2013). For this research, the authors do not intend to evaluate the acceptance or effectiveness of the platform for learning or its specific use by users during the test months, which could be determined using surveys after the pilots or based on student grades, but the authors do intend to explore user behavior and usage

patterns to obtain deeper knowledge and explore the value of this knowledge in decision-making processes about the learning process in this Virtual World.
The tests involved 75 users, and data were collected on various aspects that might indicate usage characteristics, behavior and engagement indicators in the Virtual World. These 75 users were classified as follows:

- One system administrator. This user was not included in the collection and analysis of data because the number of actions performed in the 3D environment and the time used in the system introduced too much noise in the data set.
- Ten teachers. These teachers were labeled differently in the analysis to identify potential differences in behavior between teachers and students.
- Sixty-five students. Students in the various knowledge areas described above (Pharmacy, Biology and Biotechnology, Law, Humanities and Chemistry) volunteered to try this new learning environment at the University. Because the testing stage was more focused on knowledge about use and engagement, the analysis did not differentiate among the knowledge areas of the students.

Data concerning four of the key features and options in Educational Virtual Worlds (Cruz, Therón, Pizarro, & García-Peñalvo, 2013) were retrieved: voice chat-based features, text communication-based features (between users, messages between objects, etc.), session information (time in each, total number, average time, etc.), and movement inside the Virtual Worlds (between different lands, islands, etc.). These features were chosen for analysis because they provide a range of information about users' desires, interests, usage patterns, or habits in both Educational Virtual Worlds and other eLearning Systems (Beer et al., 2010; Dalgarno & Lee, 2010; Fetscherin & Lattemann, 2008):

- Voice API calls: Data on the number of times the 3D characters used the voice service within the Virtual World. These data provide information about communication only between users and can reveal certain social characteristics (Borner & Lin, 2001; Borner & Penumarthy, 2003).
- Messages with objects and users: Data on the number of times that 3D characters made use of private messages to each other or received information from the different objects in the virtual environment. These data provide information on both social use, such as the 3D scanning environment, and the use of educational resources (Schmidt & Laffey, 2012).
- Sessions: Data about the time spent by users within the Virtual World and its distribution in time, number of sessions, average session duration, etc.
- Movements: Data about the movement patterns of users. The number of movements may be an indicator of engagement, while movement to a specific territory may indicate interest in that territory (Cruz-Benito, Therón, García-Peñalvo, & Pizarro Lucas, 2013; De Freitas, 2006).

Data about other indicators, such as clicks in 3D contents or objects, text discussion themes, changes in avatar appearance, and profile modifications and descriptions, were discarded because they are more difficult to measure in an initial exploratory analysis of this type of platform.

2.3 Identification and validation of patterns

To identify patterns, understand how users use the environment, and determine their involvement and engagement with the platform, an exploratory analysis of the data was performed. A detailed exploration of the data was performed to reveal trends and relevant issues in the collected set of interaction-evidence. A panel of experts in the creation and use of learning scenarios in Educational Virtual Worlds was interviewed to obtain some general heuristic rules and their application and validation or rejection based on trends and relevant data.

The expert panel comprised five professors from the University of Salamanca, none of whom was an author of this paper or interfered actively in the realization of this research study beyond the discussion of patterns or behaviors. The experts had five years of experience in the definition and use of Educational Virtual Worlds in a Higher Education context, specifically in Second Life, which was launched in 2003, and thus their experience encompasses nearly half the time this system has been available on the market

Individual interviews were conducted with the members of the expert panel to achieve two goals: first, to establish some heuristic rules and general occurrences observed in their experience; second, to review each of the most representative patterns and rules obtained in the exploratory analysis. Conclusions and perspectives were also collected as part of the accomplishment of these goals.

This part of the section will describe and explain the general heuristic rules disclosed by the experts; these heuristic rules will help the authors explain the patterns and association rules between the measures. Among these general heuristic rules, the following can be highlighted:

1. The experts first insisted that there are different types of users that can be highly differentiated in terms of ease of use and engagement toward these platforms. Users (students or teachers) can be classified into three groups: those who reject this technology, those who find it difficult to use but are able to employ it simply but effectively with practice, and those that operate seamlessly in these environments, display strong interest in it and are intensively engaged with the virtual environment. The experts also emphasized that this classification is often not reflected in the scores of activities conducted within the Educational Virtual World, although groups that use these tools more proficiently and with greater engagement often used more resources and achieved better results.
2. Consistent with the previous heuristic rule, the experts emphasized *that users who spent more time on the learning scenarios (not in the Educational Virtual World, but in a particular learning scenario) tended to perform better in the learning activities*.
3. Moreover, the experts stated that *users without extensive experience in the use of tools and 3D (Virtual Worlds, video games, etc.) often refuse or are reluctant to use the virtual environment, occasionally even resulting in dropouts from the activity scenarios*. These users may also underperform because of their non-acceptance of technology.
4. The types of features and resources used depend on the learning activity performed. Experts agreed that *users use more 3D objects, text-based chat or voice chat*

depending on the learning practice; for example, in practices with virtual face-to-face meetings, user tend to use chat (both text- or voice-based) more, while in autonomous learning activities, users tend to interact more with virtual objects.

3. Theory and Calculation

3.1 Theory

This section describes some key aspects used in the article to inform the reader of the concepts used for exploratory analysis. These concepts are about engagement and behavior patterns, their relationship with the learning ecosystems, and their relevance in Educational Virtual Worlds. The second part of this section describes the used by the authors to perform the analysis and achieve the results.

As explained previously, engagement is a key factor in the successful adoption of a learning ecosystem by students, but one question must first be addressed: *What is engagement?*
Many authors have attempted to define engagement in the context of online learning environments (Beer et al., 2010; Bulger, Mayer, Almeroth, & Blau, 2008; Krause & Coates, 2008; Zhao & Kuh, 2004) using different factors, either in terms of effectiveness, the element usage rate, platform acceptance and use, motivation, etc.
To clarify the engagement concept in a simple definition, it can be asserted that *engagement is related to the involvement of students in their learning process or tools*. This involvement is evident in many ways, such as the time users spend learning, the number of activities undertaken, the relationship between results quantity and quality, etc. (Stovall, 2003).
In general, engagement factors are those characteristics that can induce a special or superior interest in its learning cycle (Gómez Aguilar, García-Peñalvo, & Therón, 2014). In the case of Virtual Worlds, for example, engagement features are those factors related to the interactions of users with both the 3D environment and other users that increase use and acceptance of the environment, thereby achieving better final results in the learning process.

As explained previously, the measured factors for this research work are the following: data about voice API calls, data about messages with objects and users, data about sessions, and data about movements. As stated previously by others, the interaction between the user and the environment is more reliable in Virtual Worlds than in other eLearning ecosystems because users identify their digital identity (the 3D character or avatar) as a replica or reliable representation of themselves (De Freitas, 2006; De Freitas & Neumann, 2009).
These factors all involve interaction with the environment and other users or the amount of time and the rate of use of the system or specific features of it.

Measuring the engagement factors reveals common factors, characteristics, behavior or usage patterns among users. As many authors have noted (Bulger et al., 2008; Dalgarno & Lee, 2010; Davis, 1989; De Freitas & Neumann, 2009; Shen & Eder, 2009; Virvou & Katsionis, 2008; Wojciechowski & Cellary, 2013), these usage patterns and evidence of engagement can help explain how users engage with certain learning tools, enabling improvements or corrections of aspects that are not being used correctly, preventing dropout, and helping users to reach their learning goals (Gómez Aguilar, Therón, & García-Peñalvo, 2009).

3.2 Calculation

As mentioned in subsection 2.1, the exploratory analysis of data retrieved from the Virtual World was performed using Microsoft Excel software. Once the proposed factors for engagement and usage were measured, they were counted, sorted, and clustered in groups of users, relating each user with performance and the engagement factor measures. These clusters and rules are obtained by counting and measuring the factors and joining primary factors with other related factors to obtain rules such as the following: *users who spend more time in the Virtual World use text-based chat more extensively*. To determine which groups or clusters of occurrences may constitute a behavior pattern, the following rule was used.

A behavior pattern can be established by observing significant overlap in the interaction and use measures by at least 15% of the users of the system. Significant overlap is established when a difference or deviation occurs in the measurements between users and other of at least 30% of the measurement.

This rule is based on considerations of association rules presented by Agrawal and others in 1993 (Agrawal, Imieliński, & Swami, 1993). This research work, one of the main references in the field of association rules, explains the protocol for selecting relevant item sets and items within a dataset to create associations and rules regarding the data of a system. To select these relevant data, values and factors are used to discover rules; here, the authors defined 15% of the population (minimum confidence of the association rule) as a substantial number of users with similarities in any measure and with a 30% difference from the same issue measured in remaining population (for example, a difference of 30% of the total movements between some users and the others); this 30% is the support of the rule and is based on the support difference between the minimum confidence (at least 15% of the population) of the rule and the other part of the population (Mannila & Toivonen, 1996, 1997; Srikant, Vu, & Agrawal, 1997).

For cases with less disparity of factors and values (the presented case study has a small number of users with substantial disparity and heterogeneity in their characteristics based on different issues such as skills related to computer-interaction, age, etc.), values higher than 15 and 30% could be used to obtain greater accuracy in rule discovery, but in the present study, the authors fixed these values to validate several rules to avoid empirical heterogeneity in the observations of the users' usage of the system.

To apply this rule, the authors employed the following process with the software used to explore the data. Once the data retrieved from the Virtual World were entered in Microsoft Excel, it was classified in several columns according to the parameter measured (e.g., calls to API voice, instant messages, sessions, total time logged in, session number, and mean time), and each measure was related to the corresponding user. Once the data were organized in the spreadsheet, the authors searched for significant gaps between the data (Figure 1), such as users with a greater significant measure (30% gap) compared to the rest of the users (this gap facilitated the determination of the support of each rule that will be discovered later and must be larger than 30%). The authors subsequently compared each value and selected users corresponding to coincidence in the significant use of each resource. When the authors observed that more than 15% of the total population of the case study (15% was the minimum confidence of the future association rules proposed) exhibited coincidence in the range of the usage measure of the resources, a new pattern in the usage of the system was established, and a rule was formulated based on these data and the relationships discovered.

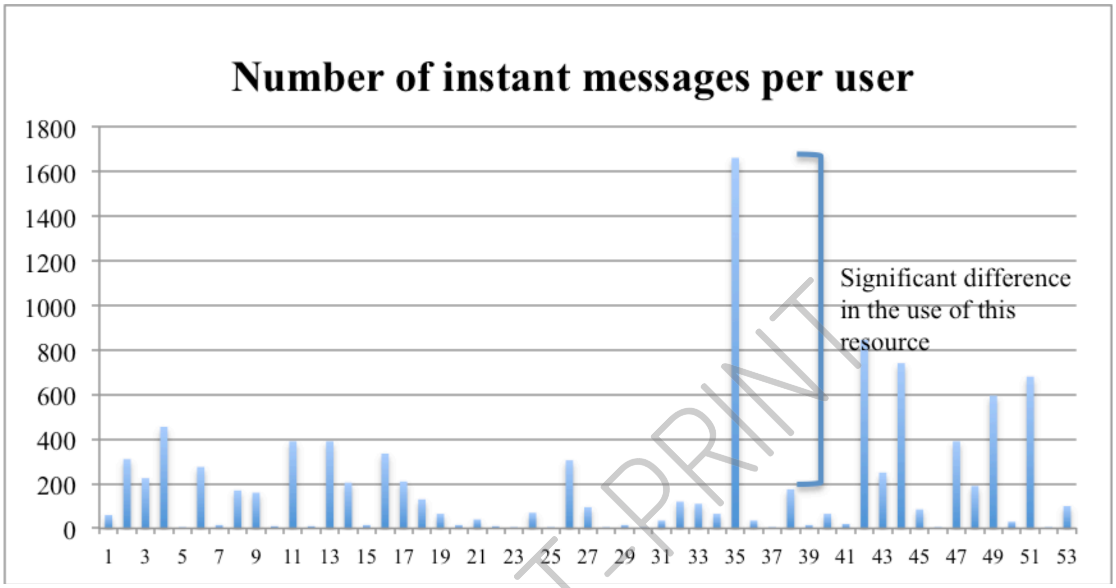


Figure 1. Example of significant differences in the use of a Virtual World resource

For example, to consider that some users have a movement rate between islands of the Virtual World which may be considered as behavior pattern, this occurrence has to occur in at least 15% of them, and the difference in the measure of the movement (in and out of different regions) the selected group from the rest of users must be at least 30%. And if the pattern is expressed as a rule, for example: *users with more time spent in the Virtual World use more the text-based chat*, both behaviors, *more time spent* and *use more the text-based chat* must to meet this rule of 15% - 30%.
The results obtained performing exploratory analysis through Excel software and applying this rule are shown in the following section (4. Results).

4. Results

The exploratory analysis revealed many usage and behavior patterns. As explained in the second part of section 3 (3.2 Calculation), the primary patterns were then linked with related patterns (secondary patterns involving the same clusters of users). Additional patterns and characteristics that were identified in the data analysis but were not sufficiently representative or were only residually present were not included in this study.

After identifying the patterns, the authors consulted through personal interviews the expert panel to determine the accuracy of each feature or interesting pattern. The patterns and rules obtained from the data analysis and the opinion of the experts were as follows:

- *Users who participate in more sessions use more resources and tools in the Educational Virtual World*. Users who logged more time in the Virtual World (as measured by number of logins and not time spent in the virtual world) often used a greater amount of resources (more interaction, voice and text messages, etc.). In particular, a relationship between the number of logins and text messages was noted. Interestingly, the number of resources used increases exponentially for users who connect often and is not a simple linear relationship; as the number of sessions and, possibly, acceptance of the 3D environment increases, users tend to perform a greater number of actions and exhibit higher ratios of actions per session. Of the 20 users with more sessions (in terms of total users and taking into account the 15% - 30% rule), 12 exhibited this usage pattern (in Figure 2, each number on the X axis is one user, and the Y axis shows the number of sessions, voice API calls or instant messages, depending on the chart), corresponding to users 2, 3, 4, 16, 34, 55, 62, 63, 64, 67, 71 and 73. As shown in Figure 2, these users occupy the top positions in the ranking of use of text-based messages and voice features in the Virtual World.

  The experts agreed that this pattern is a common feature in Virtual Worlds as well as educational and other environments. The relationship between the usage and advantage of a tool can be expressed as a logarithmic curve, indicating that users who use a tool frequently know it deeply and exploit it more. Related to the first heuristic rule, users from the third group of the rule (users who operates seamlessly in these environments), tend to make better use of the 3D environment and therefore spend more hours learning the features and options that the Virtual World offers.

- *Users who spend more time in the Virtual World tend to have more interactions with objects than with other users.* Among users who spent the most time (20 of the total users) in the Virtual World, more than half interacted more with objects inside the 3D scenarios than with other users of the experience.

  According to the experts, this pattern reflects the purely academic purpose of the Virtual World in which the data were collected. From their perspective, this pattern reveals that users who have spent more time in the Virtual World have spent more time exploring the 3D resources and solving the problems and learning challenges they entail. The experts assumed that users who use Virtual Worlds for hours are not always conducting practice or learning activities and thus often tend to explore the environments, usually interacting with the environment and the 3D objects. Without comparing it to academic performance (which could settle whether this engagement is based on academic aspects or similarity to video games), this pattern indicates that a user devotes time to the environment because he or she is interested in exploring it and improving the use of features included in the learning scenarios (second heuristic rule). The experts also indicated, as described in the fourth heuristic rule, that this goal is not always achieved because students must interact with other users in some learning activities (such as role-playing practices), and thus this result can be considered only as a behavior pattern in this Virtual World and not as a general rule applicable to any similar environment.

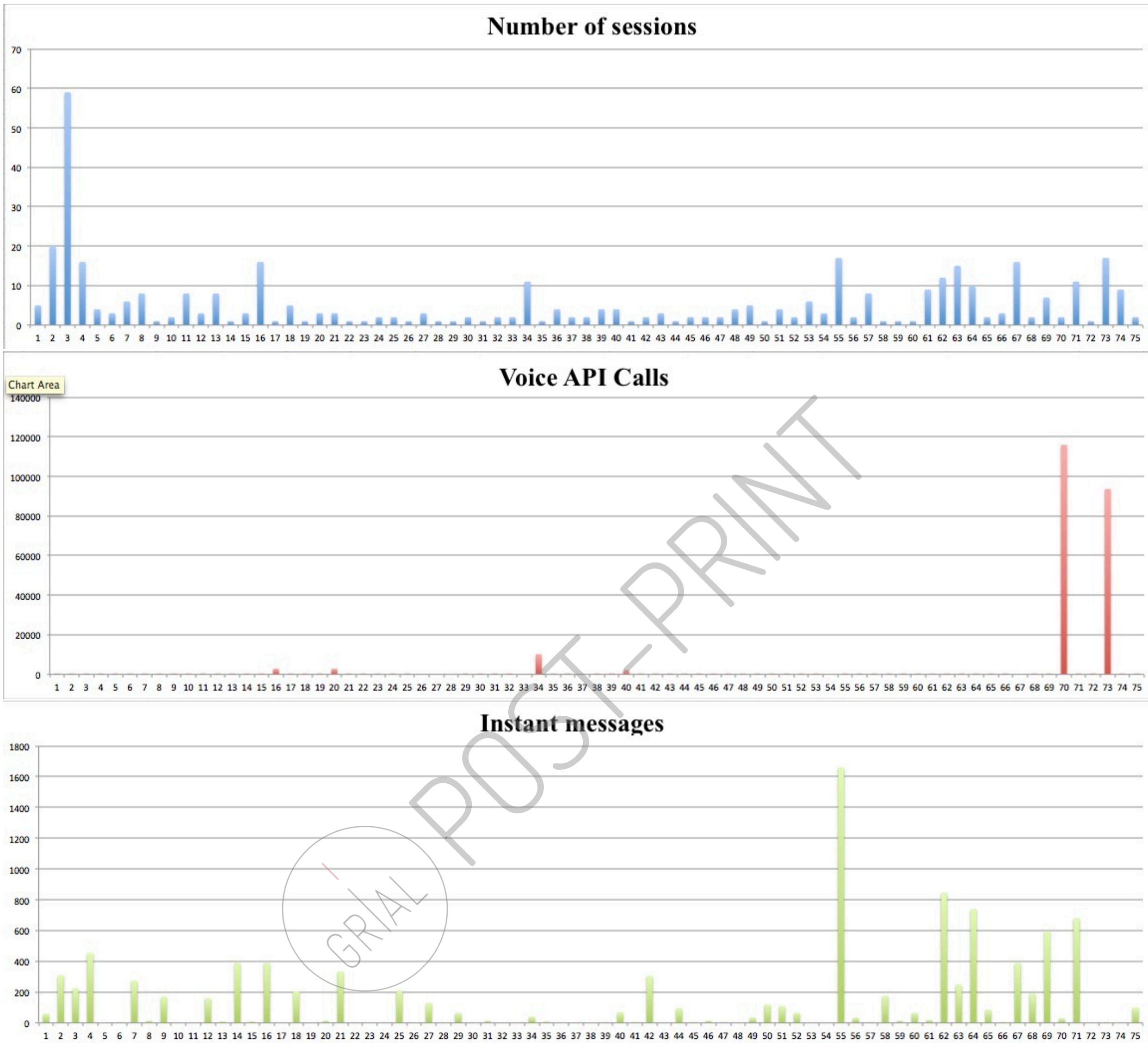


Figure 2. Charts comparing the number of sessions, use of Voice API and number of Instant messages among users.

- *Teachers use a greater amount of voice resources.* Six teachers were among the 20 users with the greatest use of the voice service within the Virtual World (the average number of Voice API calls by users and the standard deviations of these data are presented in Table 1). These six corresponded to users 3, 17, 20, 34, 70, 73. This statement is considered only a usage pattern and cannot be related to another pattern to establish a rule.

  This pattern was confirmed in the fourth heuristic rule of the expert panel. Teachers use voice resources more because they tend to talk to students in these contexts, either to engage in conversation through their virtual alter ego (without the rigidity of the teacher-pupil relationship) or because many teachers are more comfortable clicking a button and talking than writing on the keyboard at the same speed as students (who usually write more quickly). For students, this difference in the use of verbal or written communication is often not obvious because they are much more accustomed to using text-based communication in instant messaging systems and existing social networks.

Table 1. Number of Voice API calls by the six teachers who were among the most active users of voice services within the Virtual World (average number of Voice API calls = 3085.7; Standard Deviation of the data related to Voice API calls = 16979.26)

User number	*Voice API Calls*
User 3	*160*
User 17	*3513*
User 20	*2810*
User 34	*10190*
User 70	*115955*
User 73	*93625*

- *There is no relationship between the number of sessions and total time spent in the Virtual World.* None of the 20 users with the greatest number of sessions on the virtual platform were among the 20 users that spent the most time in it (Figure 3). Thus, many users connect several times for short durations, while others connect less often but for much longer durations; these two patterns are not directly related.

  The experts could not provide a clear explanation for this behavior. Two suggested that it might be related to testing experience with this Virtual World, such that many users are conducting connection tests in the virtual world, while the most skilled users performed less connection tests but remained connected for longer time periods. In summary, the experts did not perceive a regular pattern in these 3D environments, and the lack of relationship between number of sessions and time may be either a temporary or unique feature of this environment.

- *Users who refuse this technology stop using it very early.* Of the 20 users who spent the lease amount of time in this environment, 16 had spent less than 3 minutes online. In most cases, they had also used little or none of the other resources analyzed (voice, messages), as shown in Table 2.

  All experts agreed that this type of platform produces very sharp reactions among users. As indicated in the first and third heuristic rules, users with no experience with video games or 3D environments frequently have experience difficulty in moving and using the avatar in the virtual environment, causing some to leave these systems very early. Moreover, the experts also suggested that students may not have taken the voluntary testing seriously and thus conducted only a connection test without actually experiencing true use of the platform.

Table 2: Comparison of total usage time (seconds), voice API calls and instant messages among the 15 users who spent the least time in the Virtual World

Total usage time (seconds) *Average (all users) = 1459.99* *Standard deviation = 1758.9*	*Voice API calls* *Average (all users) = 3085.67* *Standard deviation = 16979.26*	*Instant messages* *Average (all users) = 136.33* *Standard deviation = 255.56*
5	*0*	*0*
8	*0*	*0*
32	*0*	*0*
43	*2810*	*15*
49	*0*	*0*
68	*70*	*15*
96	*60*	*0*
115	*80*	*0*
119	*80*	*205*
125	*40*	*0*
126	*50*	*60*
131	*90*	*20*
144	*160*	*225*
156	*0*	*390*
161	*30*	*210*
168	*40*	*160*

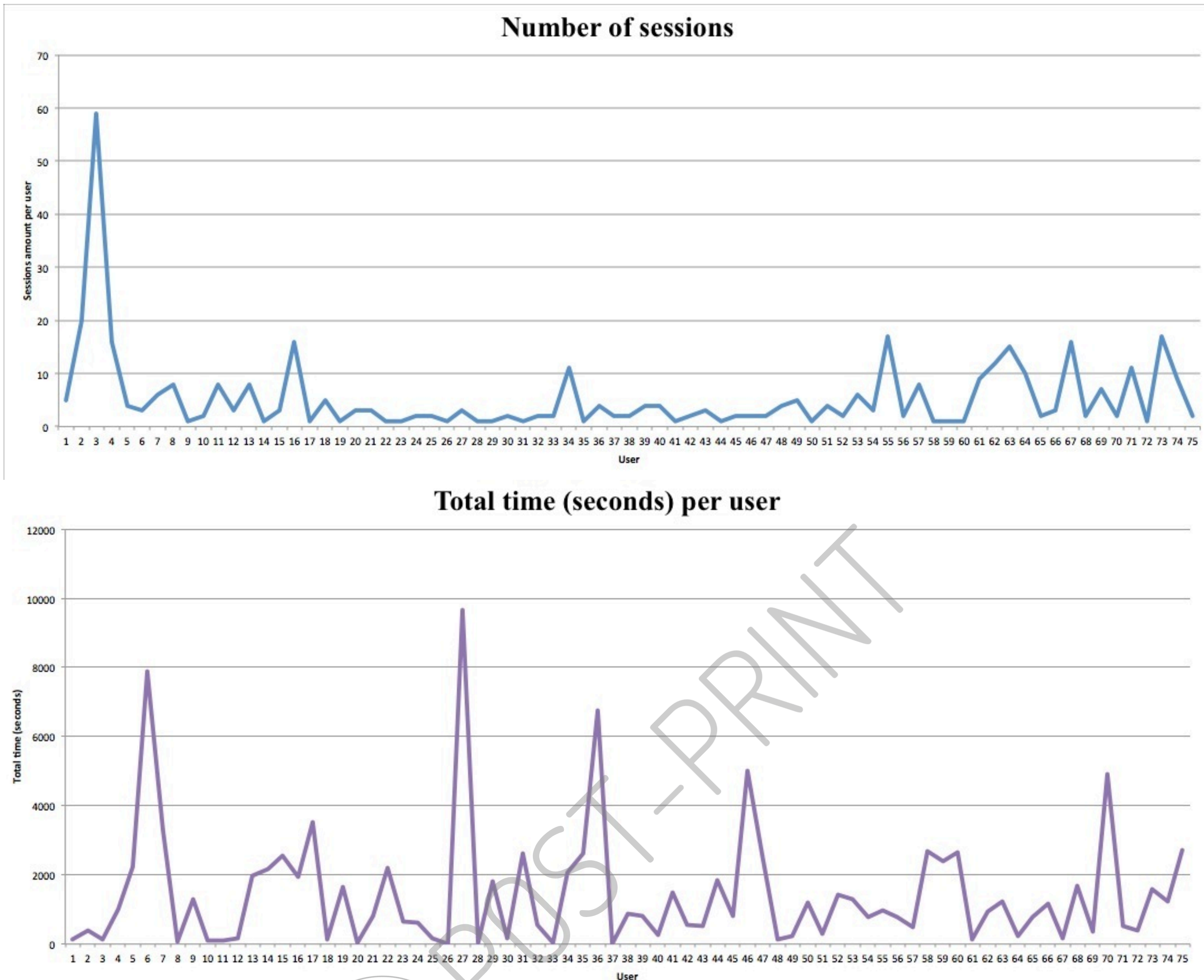


Figure 3. Comparison of the number of sessions and total usage time per user.

In summary, and in terms of engagement, the experts group believed the most relevant patterns were the following: *users with more sessions use more resources* and *users who spend more time in the virtual world tend to interact more with objects than with other users.* These two patterns represent situations of pure engagement either in the platform or in the learning content.

In addition, according to the panel, a greater in-depth analysis of object usage or voice-based chat usage might reveal if the engagement in these patterns is motivated by social relationships, the learning system, or purely entertainment based on its visual similarity to a videogame.

For the pattern *users who refuse this technology stop using it very early*, there is no engagement with the system. This pattern could be very useful for identifying reasons for a lack of interest in the virtual environment among a group of relevant users and plan actions to avoid the dropout of these users.

Moreover, experts associated the pattern *Teachers use a greater amount of voice resources* with the teachers' ability to use the voice-based chat because it is simpler and faster for teachers to talk than write. The experts indicated that this usage of voice-based chat reflects the teachers' desire to communicate actively with students for, at least in part, academic reasons; such high use suggests that teachers are comfortable with the tool and have some engagement with it.

5. Discussion

Is this type of study useful? Does this study differ from similar studies of other systems used for learning? As many authors agree, among the community that engages in the formal use of Educational Virtual Worlds and retrieves information from these systems, many researchers and early adopters of this type of environment for education have recognized a gap between the events inside the virtual world and the teachers' knowledge of these events (Atkinson, 2008; García-Peñalvo, Cruz-Benito, Maderuelo, Pérez-Blanco, & Martín-Suárez, 2014; Warburton, 2009). This gap cannot be addressed easily. The data and evidence retrieved from other types of environments are very different than those that can be collected from Virtual Worlds; these differences may be due to the nature of the student fingerprint, which requires new approaches to understand user interactions, or to the private nature of some of these environments (Baker, Wentz, & Woods, 2009; Cruz et al., 2013; Varvello, Ferrari, Biersack, & Diot, 2011; Warburton, 2009). This type of study is useful because obtaining knowledge about the users' usage and interaction with the Virtual World and learning through the contents and resources presented within a Virtual World will help teachers and managers of the Educational Virtual Worlds better plan and design the deployment of educational content and resources inside the 3D environment, enhancing the personal experience and learning process for students.

Accordingly, how can managers use this knowledge to improve the user engagement and the utility of the Virtual World? Based on the cases and rules described above, there are several rules (those based on usage behavior) that managers of the virtual environment could follow to help users make better use of the environment and meet their learning goals.

For example, in the case of the rule *Users who refuse this technology stop using it very early,* managers could plan actions, reminders, classroom activities to help students, and, if this proportion of students is very large, plan introductory workshops about this technology. Knowledge about early dropout from this technology and environments allows managers to remain vigilant about users' engagement indicators and act to avoid dropout among users meeting this rule.

In the case of the rule *Teachers use a greater amount of voice resources,* the managers can focus their attention on voice chat usage indicators to identify teachers who are not using voice chat and may be experiencing technical problems or are unsure how to use it correctly.

In the case of rules reflecting engagement (for example, *Users with more sessions use more resources*; *Users who spend more time in the virtual world tend to interact more with objects than with other users*), these indicators could serve to measure interest in the different features available in the 3D environment to allow managers, teachers, designers and developers of the virtual world to successfully design new learning scenarios that are more interesting to students or the general users of the Educational Virtual World. These indicators of engagement and types of user profiles can also allow managers, designers, etc.

to design activities specifically for each type of user. Thus, different features and scenarios could be offered to users who use and perceive the environment differently to facilitate learning goal achievement.

6. Conclusions

Usage behaviors and engagement indicators represent an appropriate approach with which managers of the Educational Virtual Worlds can determine how users use the online features available, which tools have greater acceptance, and which tools are more useful for users. This knowledge allows managers to make decisions and plan actions about the eLearning system; these decisions and actions could cover several aspects, such as preventing dropout from learning activities, promoting interesting features to increase user engagement and engage more users, create learning scenarios adapted to users' detected desires, etc.

This paper demonstrates that behavior patterns and engagement factors can be detected and measured in a simple manner via exploratory analysis, followed by an evaluation of the validity of the patterns and measures by a panel of experts. The experts agreed with the results of the analysis, with the exception of *There is no relationship between the number of sessions and total time spent in the Virtual World*; the experts did not consider this a common pattern among other Educational Virtual Worlds and 3D-based learning scenarios. The other patterns and measures are related to engagement with available features and content and may reflect engagement with the Virtual World as a technology itself or the available content (learning content). The difference between engagement based on technology or based on content may be relevant to decision-making and planning by managers to enhance the Educational Virtual World as an eLearning tool or enhance the educational content available.

Ideally, in future work, this detection and measurement could be performed automatically and in a more standardized manner, such as by using data mining algorithms (for example, association rules algorithms) or deploying a framework to retrieve, organize, analyze and present information transparently to managers and other staff related to Educational Virtual Worlds to allow them to exploit this knowledge in real time to plan relevant decisions and actions.

References:

Agrawal, R., Imieliński, T., & Swami, A. (1993). *Mining association rules between sets of items in large databases.* Paper presented at the ACM SIGMOD Record.

Atkinson, T. (2008). Inside Linden Lab: Second Life [TM] for Educators. *TechTrends: Linking Research and Practice to Improve Learning, 52*(3), 16-18. doi: 10.1007/s11528-008-0147-8

Baker, S. C., Wentz, R. K., & Woods, M. M. (2009). Using virtual worlds in education: Second Life® as an educational tool. *Teaching of Psychology, 36*(1), 59-64.

Beer, C., Clark, K., & Jones, D. (2010). Indicators of engagement. *Proceedings ascilite Sydney*.

Borner, K., & Lin, Y. C. (2001). *Visualizing chat log data collected in 3-D virtual worlds.* Paper presented at the Information Visualisation, 2001. Proceedings. Fifth International Conference on.

Borner, K., & Penumarthy, S. (2003). Social diffusion patterns in three-dimensional virtual worlds. *Information Visualization, 2*(3), 182-198.

Bulger, M. E., Mayer, R. E., Almeroth, K. C., & Blau, S. D. (2008). Measuring learner engagement in computer-equipped college classrooms. *Journal of Educational Multimedia and Hypermedia, 17*(2), 129-143.

Cruz, J., Therón, R., Pizarro, E., & García-Peñalvo, F. J. (2013). Knowledge Discovery in Virtual Worlds Usage Data: approaching Web Mining concepts to 3D Virtual Environments. *Proceedings Fourth International Workshop on Knowledge Discovery, Knowledge Management and Decision Support (Eureka-2013)*.

Cruz-Benito, J., Therón, R., García-Peñalvo, F. J., & Pizarro Lucas, E. (2013). *Analyzing users' movements in virtual worlds: discovering engagement and use patterns*. Paper presented at the Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality, Salamanca, Spain.

Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology, 41*(1), 10-32.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.

De Freitas, S. (2006). Learning in immersive worlds. *London: Joint Information Systems Committee*.

De Freitas, S., & Neumann, T. (2009). The use of 'exploratory learning'for supporting immersive learning in virtual environments. *Computers & Education, 52*(2), 343-352.

Fetscherin, M., & Lattemann, C. (2008). User acceptance of virtual worlds. *Journal of Electronic Commerce Research, 9*(3), 231-242.

García-Peñalvo, F. J., Cruz-Benito, J., Maderuelo, C., Pérez-Blanco, J. S., & Martín-Suárez, A. (2014). Usalpharma: A Cloud-Based Architecture to Support Quality Assurance Training Processes in Health Area Using Virtual Worlds. *The Scientific World Journal, 2014*.

Gómez Aguilar, D. A., García-Peñalvo, F. J., & Therón, R. (2014). Analítica Visual en eLearning. *El Profesional de la Información, 23*(3), 233-242.

Gómez Aguilar, D. A., Therón, R., & García-Peñalvo, F. J. (2009). Semantic Spiral Timelines Used as Support for e-Learning. *J. ucs, 15*(7), 1526-1545.

Krause, K. (2005). Understanding and promoting student engagement in university learning communities. *Paper presented as keynote address: Engaged, Inert or Otherwise Occupied*.

Krause, K., & Coates, H. (2008). Students' engagement in first-year university. *Assessment & Evaluation in Higher Education, 33*(5), 493-505.

Lucas, E. P., Cruz–Benito, J., & Gonzalo, O. G. (2013). USALSIM: learning, professional practices and employability in a 3D virtual world. *International Journal of Technology Enhanced Learning, 5*(3), 307-321.

Mannila, H., & Toivonen, H. (1996). *Discovering Generalized Episodes Using Minimal Occurrences.* Paper presented at the KDD.

Mannila, H., & Toivonen, H. (1997). Levelwise search and borders of theories in knowledge discovery. *Data mining and knowledge discovery, 1*(3), 241-258.

Messinger, P. R., Stroulia, E., Lyons, K., Bone, M., Niu, R. H., Smirnov, K., & Perelgut, S. (2009). Virtual worlds — past, present, and future: New directions in social computing. *Decision support systems, 47*(3), 204-228. doi: http://dx.doi.org/10.1016/j.dss.2009.02.014

OpenSimulator. (2014). OpenSim Web Page. Retrieved 1/02/2014, from http://opensimulator.org

Schmidt, M., & Laffey, J. (2012). *Visualizing Behavioral Data from a 3D Virtual Learning Environment: A Preliminary Study.* Paper presented at the System Science (HICSS), 2012 45th Hawaii International Conference on.

Shen, J., & Eder, L. B. (2009). Intentions to Use Virtual Worlds for Education. *Journal of Information Systems Education, 20*(2).

Srikant, R., Vu, Q., & Agrawal, R. (1997). *Mining Association Rules with Item Constraints.* Paper presented at the KDD.

Stovall, I. (2003). Engagement and Online Learning. *UIS Community of Practice for E-Learning*.

Varvello, M., Ferrari, S., Biersack, E., & Diot, C. (2011). Exploring second life. *IEEE/ACM Transactions on Networking (TON), 19*(1), 80-91.

Virvou, M., & Katsionis, G. (2008). On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE. *Computers & Education, 50*(1), 154-178.

Warburton, S. (2009). Second Life in higher education: Assessing the potential for and the barriers to deploying virtual worlds in learning and teaching. *British Journal of Educational Technology, 40*(3), 414-426.

Wojciechowski, R., & Cellary, W. (2013). Evaluation of learners' attitude toward learning in ARIES augmented reality environments. *Computers & Education*.

Zhao, C.-M., & Kuh, G. D. (2004). Adding value: Learning communities and student engagement. *Research in Higher Education, 45*(2), 115-138.

6.3 Appendix C. Monitoring and feedback of Learning Processes in Virtual Worlds through analytics architectures: a real case

Monitoring and feedback of Learning Processes in Virtual Worlds through analytics architectures: a real case

Juan Cruz-Benito, Roberto Therón, Francisco J. García-Peñalvo
GRIAL Research Group
Department of Computers and Automatics
University of Salamanca
Salamanca, Spain
{juancb, theron, fgarcia}@usal.es

Cristina Maderuelo, Jonás Samuel Pérez-Blanco, Hinojal Zazo, Ana Martín-Suárez
Department of Pharmacy and Pharmaceutical Technology
University of Salamanca
Salamanca, Spain
{cmaderuelo, jsperez, hinojal, amasu}@usal.es

Abstract— **The application of techniques related to Learning Analytics, including the recovery and processing of data and evidences related with the learning tasks, might result capital to improve educational processes that occurs within the virtual worlds. This paper aims to show a real example of application of a learning analytics framework which supports a real case of practical learning in a 3D virtual environment, analyzing in depth the problems that arises therefrom. As remarkable outcomes, this paper shows a whole system that comprises several tools, which make possible the automatization of a learning process inside a virtual world with the subject of quality assurance in pharmacy laboratories.**

Learning Analytics; Virtual Worlds; Second Life; Framework for analytics and learning support; Learning Evidences; Monitoring

I. INTRODUCTION

Since the beginning of the use of Virtual Worlds in educational contexts, several authors have pointed out the problems related to know what is happening within the 3D environment about users behaviour, use of virtual world, etc [1] [2] [3]. Among the goals of this knowledge of what happens in virtual worlds, highlights the possibility of evaluate *offline* (outside the Virtual World) what happens *online* (inside the Virtual World). Achieving this goal, teachers related with Educational Virtual Worlds could know completely the learning process that takes place within it, so they could measure the learners evolution, interest, contents learned, results they get, or even predict anormal behaviour, dropout of students, etc.

From students point of view, the data retrieved from their usage of virtual world, could help them to improve their learning process, for example, if there is a system integrated with virtual world that guides them to achieve learning goals [4], discovering new training process, warning them if they fails, scoring their progress and achievements in educational activities, etc. like a virtual *master* who help them improve. The presence of a supervisor system which analyzes what happens in the virtual world, and can give results and answers to *in-world* events, can facilitate the implementation of training scenarios with a closed number of paths or solutions, so that the system itself could assess achievement of the different phases or solutions. Thereby, the digital supervisor could help automate the educational process, or professional training that occurs within the virtual world, facilitating independent learning and repetition of closed cycles of knowledge acquisition by students.

This paper aims to describe a real case application of a system that analyzes what happens inside a closed 3D environment (a virtual laboratory), and report data about it to students who are in the Virtual World, and teachers who are out of it. For this purpose, the paper include the following sections: the first one serves as an introduction to the common problem to deal with, the second, called *USALPHARMA Lab: Case study Description*, is about the real context in which it has applied one of these analysis platforms, the third section (*Architecture*) describes the overall process of retrieval and analysis of data including its components, the fourth section (*Data Retrieval from Virtual World*) describes how are extracted data from 3D environment to be stored and processed after. The fifth section, called *Data Storage and Analysis* is about how the data is stored and treated to meet the queries and analysis tasks of users. Meanwhile, the sixth, *Feedback to Users in Virtual World*, discusses how students access from the Virtual World to the reports about their learning and progress. The seventh section (*Feedback to Teachers with a Web Client*) describes the query tool provided to teachers in order to consult data of students and their training process. Finally, it includes a section on conclusions, which also presents possible future work.

II. USALPHARMA LAB: CASE STUDY DESCRIPTION

USALPHARMA Lab is a virtual laboratory located in USALPHARMA island in Second Life, that simulates the installations, equipment, documentation and tools like a real quality control laboratory of the pharmaceutical industry that complies with GLP (Good Laboratory Practice) to such effects [5]. These types of installations in the real world are very expensive and unusual in Universities. For this reason, some members of development team built this virtual laboratory, so the students could train in a realistic simulation of a pharmaceutical laboratory. USALPHARMA Lab has been used for training of pharmacy postgraduates in Quality Assurance since 2011. To carry out these practices, both avatars of teachers and students, meet in the virtual laboratory and the teacher guides and evaluates the student during the activity. But, for responsible teachers of activity it would be ideal that the students could access the laboratory whenever they wanted using it as self-training tool. Also it would be ideal for teachers, so they could assess *offline* the students training process in Quality Assurance for pharmaceutical environments.

For these purposes, possible requirements for data collection and analysis tasks desired by the teachers responsible for training through the Virtual World were collected, so that would serve as the basis for the implementation of the architecture analysis. As a summary of these requirements, it is possible to highlight the following:

1) Evidences retrieval about interaction with certain objects (checkpoints) within the Virtual World.

2) Right description about user interaction with training environment.

3) Possibility of analyzing about data retrieved from virtual world in two ways:

a) Students could know information about their training from inside the Virtual World.

b) An analytics tool for teachers that would allow to get information about students, including filter features such as filtering by time, objects name, usernames, etc.

Following these requirements, has been designed system that performs tasks, and provides a framework for the integration of this type of software entities related to the recovery and analysis of data related to educational settings. One approach to this framework is described in the following sections by applying it to the specific problem of the virtual laboratory.

III. ARCHITECTURE

This analysis framework is divided into several components or layers (as show in figure 1), accordingly to their specific purpose within the overall function. Among these different functionality layers, it is possible to remark the following:

1) Evidence description layer: this layer is responsible for collecting evidence derived learning user interaction with the virtual environment (forming a digital footprint about interaction and events related with learning process). It collects all the data related to the interaction, including the moment in time, the user involved, as well as others relating to concrete action (action verb) or the location inside the 3D environment. To describe the interaction, it uses an adaptation of RDF protocol [6] (by using a description structure with subject, verb and predicate), including in the predicate the extra information about the environment where it is applied.

2) Collector layer: this component of the framework is responsible for collecting the information sent by the evidence description layer, and process it in order to store this information into the data persistence layer.

3) Storage layer: stores information so as to be always available for any component that may need it. It keeps the data properly stored to facilitate their treatment and recovery.

4) Analysis layer: performs common maintenance tasks and specific data analysis tasks to discover knowledge in the raw data stored in persistence layer. This layer of functionality, for example, is responsible for compacting and map information inside the database and to accomplish stadistical or data mining taks as necessary.

5) Presentation layer: is responsible for retrieving the information requested by the user or any application that could be integrated with this architecture and present it in a proper way for each type of request.

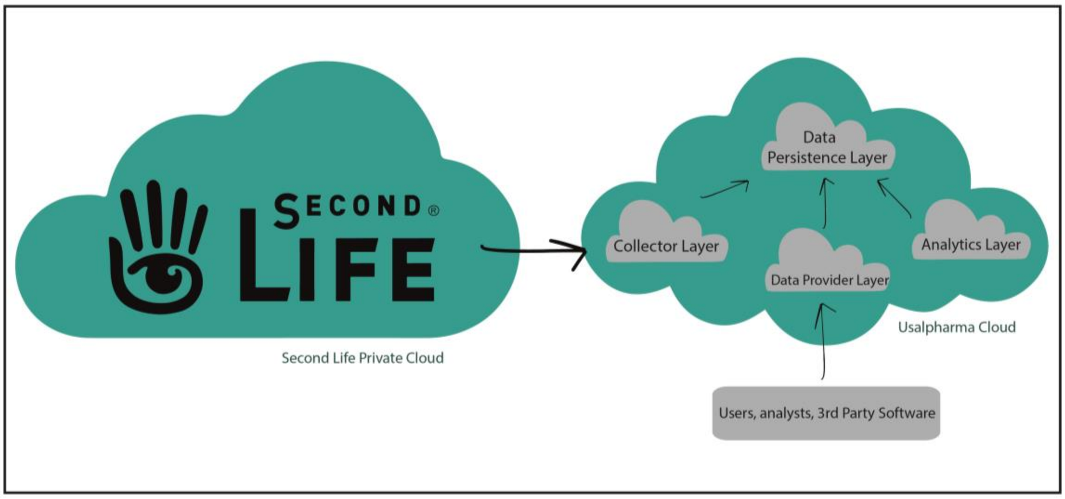


Figure 1: Cloud Architecture divided into functionality layers [7]

In our approach [7], the users are at the center of the architecture, since their interaction, both in the virtual world as the data presentation layer, leads the operation of the software components.

For example, every action that occurs into the Virtual World originated by the user interaction, triggers a series of events in the architecture: the user clicks on an 3D object, the object sends a request to a web server including data interaction, this server sets the information of the request and sends it to NoSQL data base, which stores and processes information, performing data compaction actions if necessary. The following sections will develop this cycle, giving more information about each step.

IV. Data Retrieval from Virtual World

Data recovery from the Virtual World is a fundamental task in the whole process, because if not done correctly, information that may be relevant to the analysis can be lost. Regarding the recovery of data and evidence on the interaction in the virtual world, two major questions arise: How to describe the interaction in a standardized way? and How should perform the information communication to the collecting layer?

As previously mentioned, the first question on how to describe the evidence of learning or training, can be answered through the use of a standard method for description. In the case presented, the use of RDF resource description standard can suit, modifying it in such a way that fulfills the purpose of describing user-interaction virtual environment. This modification or adaptation includes information about the user performing the action, which takes concrete action (touching, sitting, etc.), which object get the action, the time at the action is performed (time stamp), as well as other information on the virtual location where the action was performed (the laboratory, or any particular room inside it if necessary to specify). Thus, the analytics framework could use a rich data for better and powerful analysis [8], as another approaches to analysis of learning activity streams pose [9] [10] [11].

Figure 2. Example of dialog box in Second Life

Similarly, the second question can also be answered with a concept previously discussed: the user is the originator of workflows in this framework for analysis and support of the educational process in the Virtual World. For data retrieval from the architecture for analysis and support, serveral sollutions could raise, from polling techniques to data recovery driven by events. In this case study, building on the concept that the user is the originator of workflows, it is possible to raise data recovery driven by events. The user generates an event (the interaction itself) to an object in the virtual environment, and this object reacts in two ways, one responsive to the user on the interaction, and the other is sending an *interaction notification* to the data collector layer included the proposed framework. In the case study of this paper, this notification is done via an HTTP request to a web server, which collects the request parameters and prepares them to be stored in the storage layer. Depending on the activity of users in the Virtual World, it is possible that the web server has to endure considerable work load or activity peaks, so it is necessary to use appropriate technology in this regard. The web server must therefore be able to respond quickly, besides being able to be robust and scalable, even giving the possibility to be scalable on demand (as is done in cloud production environments). In this case study, it has been used a system developed under the Python language and Django framework, with an implementation that allows each of its MVC model layers could be divided into multiple virtual machines, being able thus to give high performance. Data Storage and Analysis

The data persistence layer is another key part in the architecture for analysis and support, as it stores all raw data and knowledge and information extracted from these data. In addition to storing data and knowledge, this persistence layer should be able to return results in an efficient way to queries that could be performed.

To do this, and following the same philosophy as in the previous case, has been developed the persistence layer using a storage technology designed to treat large volumes of data and the possible heterogeneity of them. Based on the advantages in this respect that provides the NoSQL databases [12] [13], has been chosen one of them, MongoDB. Due to its organization documents and subdivision into different functionality levels, and features for load balancing and treatment of huge datasets [14], MongoDB looks good for this case, even to extend the application of this analysis architecture over anothes study cases involving other datatypes. Thus, in each document that composes this type of database, are stored evidences of interaction; this kind of organization made this database be able even vary the structure of these documents in the future, allowing the use of the old documents without modifying them if it is not necessary.

In order to get more efficient results this database allow to perform maintain operations over the data, such as map, reduce, count, etc. If the system detects more common queries, and data common to many queries, it could use these features and operations to anticipate the user requirements and resolve them in a easier and faster way. For example, with these features, the storage and analytics layers could count achievements, number of users, and general statistical data scheduling them in order to have them in all moments, without perform the analytics tasks each time user requires it. this type of features and tools help to get simpler, and less load the database, and thus, get better performance in queries. As examples of automatic analysis tasks, could be posed the following: user count, interaction count, 3D objects more used, common checkpoints not achieved by students, achievements get by student per day, evolution in achievements per students over the time, etc.

Figure 3. View of web tool for teachers

V. Feedback to Users in Virtual World

Among the requirements proposed by the responsible team of 3D educative environment, there is one about feedback to students inside the Virtual World. This requirement is intended to get agile the training process in Quality Assurance, because, if the students could repeat the practice alone, and when they want, they could get better results, due they could think in the practice as a self-training process, like a challenge itself, which could stimulate them.

To show progress, successes or failures to students who are within the virtual world, it is necessary to provide the means for communication between the virtual world and the framework that support the learning. Just as in the evidences collection, has been selected for the communication with the web server (data presentation layer of the framework) via HTTP requests, so that inside the Virtual World will be placed some objects that the users could touch in order to get their progress and statistics in virtual practices. This way of working prevents the user requires any type of query syntax, or auxiliary tool, and allows them to see only the statistics that teachers deem appropriate. That is, it mandates that the queries made from the virtual world where teachers are considered more beneficial to motivate and enhance the learning process.

Once the student touches the virtual object, it establishes a connection with the server, requiring a set of data (progress in the current session, checkpoints achieved in this session compared to the previous practice session, etc). The object receives via HTTP the specific data as result of the query and processes it, displaying the student the result through a dialog box of Virtual World. Due to the nature of these types of dialog boxes (Figure 2), and communication that can have an object with a user within a virtual world, information can only be presented in two forms, or plain text, or by calling a render web that can display graphics. The problem with calling the web render, is that the user loses the 3D environment while viewing the data context, as this type of web windows occupy a great part of the screen, and can cause the user to lose the feeling for a while to be immersed in three-dimensional scene. Therefore, usually the information will be in plain text.

Figure 4. View of checkpoints achieved by a student

VI. FEEDBACK TO TEACHERS WITH WEB CLIENT

Moreover, teachers included a requirement to have some kind of tool that would allow them to consult the data retrieved from the virtual world, so that it could help them to facing the assessment of the virtual activities and self-training. For this purpose, has been developed a web interface that teachers can use to build more complex queries to the data stored in the database. This web tool (figure 3) includes a form with a set of filtering and search options, among which are: user name based queries, queries with name objects in the virtual world, filters as by specific dates interaction or date ranges, or simply other filter to show whether students have achieved the control points considered most important in the activity (figure 4). Also, teachers have a another filters for pure statistical data; part of them destinated to understand data about overall students using the virtual environment, as anothers designed to meet the estimated use per student, so that they can perform an exploratory analysis to understand the use of the tool in a global way and not to search specific use cases. With all these search parameters, teachers for example, could know if students have discovered all checkpoints inside the Virtual World, if they have used more or less the learning environment, if they logged in with other students at the same time. The combination of these search filters allows teachers to know all relevant elements of the educational use of virtual world and see the evolution in learning and goals achieved by students.

In this case, has been implemented a text-based solution, but there are other ways to do these analysis tasks, for example, it is possible to build in the future a visual application that allows to teachers to perform data exploration through visual analytics, and receiving the results of the query in a interactive way, so they could explore the data in a wide way, joining different queries and tasks without a restrictive sintax or forms [15].

VII. CONCLUSIONS

The application of this type of frameworks and platforms to support learning processes within Virtual Worlds opens new possibilities for the application of this type of teaching resources. The possibility that teachers know faithfully what happens in 3D environments makes the interaction between students and virtual resources they can be measured and assessed as if it were a traditional resource. Furthermore, the ability of this platform to give a power to students to know their learning path, opens new ways for them to know their progress and could help to motivate them, so they doesn't only see the virtual world as a video game, but also as a virtual environment with a clearly defined educational purpose.

As for the case study itself, has been reached the goals set up along with those responsible for training, giving up a functional tools (for students, teachers and data retrieval). Although it is necessary to test in a wide way the tool, in order to get a firm conclusion about its validity as approach to real time learning support, both teachers and technical team are very satisfied with this first stage of development. In order to to begin to validate this first approach to the system, it was tested during the month of March 2014 by 9 students of the Universitary Master in Drug Evaluation and Development at the University of Salamanca. These students who tested the system showed a good welcome to the system, and a fairly high usage. Right now the teachers of the subject, are evaluating the evolution of academic performance compared to previous years. In any case, these tests and validation of the tool, will help to plan the future work to improve the analysis, and support framework itself.

REFERENCES

[1]. Virvou, M., G. Katsionis, and K. Manos, *Combining Software Games with Education: Evaluation of its Educational Effectiveness*. Journal of Educational Technology & Society, 2005. **8**(2).

[2]. Wojciechowski, R. and W. Cellary, *Evaluation of learners' attitude toward learning in ARIES augmented reality environments*. Computers & Education, 2013.

[3]. Williams, D., *The mapping principle, and a research framework for virtual worlds*. Communication Theory, 2010. **20**(4): p. 451-470.

[4]. Siemens, G. *Learning analytics: envisioning a research discipline and a domain of practice*. in *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. 2012. ACM.

[5]. Maderuelo, C., et al., *Auditing Training Practice To Postgraduate Pharmacy Students In A Virtual World*, in *19th Annual Conference of the European Association of Faculties of Pharmacy (EAFP 2013)*. 2013: Ankara, Turkey.

[6]. Klyne, G., J.J. Carroll, and B. McBride, *Resource description framework (RDF): Concepts and abstract syntax*. W3C recommendation, 2004. **10**.

[7]. García-Peñalvo, F.J., et al., *Usalpharma: A Cloud-Based Architecture to Support Quality Assurance Training Processes in Health Area Using Virtual Worlds*. The Scientific World Journal, 2014. **2014**.

[8]. Siemens, G. and P. Long, *Penetrating the fog: Analytics in learning and education*. Educause Review, 2011. **46**(5): p. 30-32.

[9]. del Blanco, A., et al. *E-Learning standards and learning analytics. Can data collection be improved by using standard data models?* in *Global Engineering Education Conference (EDUCON), 2013 IEEE*. 2013. IEEE.

[10]. Glahn, C. *Using the ADL Experience API for Mobile Learning, Sensing, Informing, Encouraging, Orchestrating*. in *Next Generation Mobile Apps, Services and Technologies (NGMAST), 2013 Seventh International Conference on*. 2013. IEEE.

[11]. Software, R. *Tin Can API*. 2014 1, April, 2014]; Available from: http://tincanapi.com/.

[12]. MongoDB, *Top 5 Considerations When Evaluating NoSQL Databases*. 2013.

[13]. Lakshman, A. and P. Malik, *Cassandra: a decentralized structured storage system*. ACM SIGOPS Operating Systems Review, 2010. **44**(2): p. 35-40.

[14]. MongoDB, *Big Data: Examples and Guidelines for the Enterprise Decision Maker*. 2013.

[15]. Aguilar, D.A.G., R. Therón, and F.J.G. Peñalvo, *Semantic Spiral Timelines Used as Support for e-Learning*. J. ucs, 2009. **15**(7): p. 1526-1545.

6.4 Appendix D. Usalpharma: A Software Architecture to Support Learning in Virtual Worlds

Cruz-Benito, J., Maderuelo, C., García-Peñalvo, F. J., Therón, R., Pérez-Blanco, J. S., Zazo Gómez, H., & Martín-Suárez, A. (2016). Usalpharma: A Software Architecture to Support Learning in Virtual Worlds. IEEE Revista Iberoamericana de Tecnologias del Aprendizaje, 11(3), 194-204. doi:10.1109/RITA.2016.2589719

Usalpharma: A Software Architecture to Support Learning in Virtual Worlds

Juan Cruz-Benito, Cristina Maderuelo, Francisco J. García-Peñalvo, Roberto Therón, Jonás Samuel Pérez-Blanco, Hinojal Zazo, Ana Martín-Suárez

***Abstract*— This paper explains how a software architecture was planned, designed and implemented to help in the support of eLearning activities within a virtual laboratory inside the Second Life Virtual World. It delves into the problems related to this kind of systems and architectures and detects the possible benefits they could provide to eLearning processes. The paper shows how the software engineering principles and procedures are applied in order to solve problems like data gathering from Virtual Worlds, data analytics of information related to interaction between a user and a 3D environment, and how they can be applied to enhance the students' learning process and teachers' assessment of this learning process. This paper also shows the software product resulting of this engineering process, as well as the outcomes from the application of the solution in a real context during two academic years involving postgraduate students and subjects of Pharmacy and Quality Assurance. To end this paper, the authors explain some considerations and knowledge retrieved after this experience, focusing on the utility from the learning point of view, innovation and possible future work that could be done to improve the solution implemented.**

***Index Terms*—Software Architectures, Virtual Worlds, Technology-Enhanced Learning, Usage analytics**

I. Introduction

ENGINEERING as a discipline is, according to [1], "the systematic application of scientific knowledge to build solutions, in an effective and economically viable way, for practical problems in the service of humanity"; moreover, in the case of software engineering, it is sometimes defined as "the application of a systematic, disciplined and quantifiable development, operation and maintenance of software; i.e. the application of engineering to software " [2].

Engineering is a sort of *Swiss Army knife* that gives those who use it (engineers) the necessary tools to solve problems by varying these tools depending on the area where engineers apply their knowledge or the nature and complexity of the problem to be solved. Still furthermore, the application of this *Swiss Army Knife* is not banned to non-engineers: Engineers use it and know better than anyone how to use it, but its use can benefit the whole society. Even more specifically, it can benefit communities of individual users or work teams regardless of their status, area of knowledge and experience, to solve different kinds of problems.

This article, entitled "Usalpharma: A Software Architecture to Support Learning in Virtual Worlds," presents how engineering has been used to solve a real problem in an educational context. It describes how a multidisciplinary team composed of profiles in the world of Computer Engineering and the area of Health Sciences (specifically the area of Pharmacy) have been able to solve a complex problem, a priori, such as monitoring student activity within the context of a Virtual World to assist in the evaluation of educational activities worldwide, through the application of technical and engineering tools (in this case software engineering) [3]. Related to the application of software engineering in a multidisciplinary context and in educational settings still in development, it should be noted that this project is neither unique nor pioneer in this type of collaboration and expansion of the application spectrum of technology solutions to service education, but it brings its vision and specific solutions within the set of the current generation of education systems enhanced by technology and current trends in technology to support processes of acquisition of knowledge and skills in heterogeneous learning environments.

To illustrate this particular case, this article is divided into the following sections (in addition to this first introductory section): the second section, *Usalpharma: Educational Environment in a Virtual World*, accurately describes the context of the problem to be solved, introducing the reader to the virtual environment where the teaching of students in the knowledge area of pharmacy is developed, helping to understand the rest of the article. The third section, *The Problem: Proposal, Objectives and Requirements*, poses the challenges faced by the multidisciplinary team, as well as the objectives and requirements that the solution must meet. The fourth section, *Solution and Product Developed*, specifies the theoretical solution that was proposed after applying an approach of software engineering to the problem, and the

Juan Cruz-Benito and Francisco J. García-Peñalvo are with the GRIAL Research Group, Computer Science Department, Research Institute for Educational Sciences, University of Salamanca, Salamanca 37008, Spain (e-mails: juancb@usal.es y fgarcia@usal.es).

Roberto Theron is with the GRIAL Research Group, Computer Science Department, and Vis-USAL Research Group, University of Salamanca, Salamanca 37008, Spain (email: theron@usal.es)

Cristina Maderuelo, Jonás Samuel Pérez-Blanco, Hinojal Zazo and Ana Martín-Suárez are with the Department of Pharmacy and Pharmaceutical Technology, University of Salamanca, Salamanca 37008, Spain (e-mails: cmaderuelo@usal.es, jsperez@usal.es, hinojal@usal.es y amasu@usal.es)

translation of this theoretical solution in a tangible software product and which can be used in a real context as presented. The fifth section, *Exploitation and Results*, presents the experience gained after applying this software solution for two academic years in the activity carried out within the context of the virtual world Second Life, as well as a number of indicators of use and usefulness of that solution. The sixth section, *Discussion*, comments on the main aspects of the work that has been carried out, trying to enlighten the reader about the satisfactory and unsatisfactory aspects of the proposed solution and the experiment carried out, always from a standpoint of reflection and awareness of the results and their significance. Finally, the seventh section, *Conclusions and Future Work,* summarizes the results in general and a series of lines that open future work on this project.

II. Usalpharma: Educational Environment in a Virtual World

Usalpharma Lab is a virtual laboratory built by the teaching innovation group Usalpharma and the Department of Pharmacy and Pharmaceutical Technology of the University of Salamanca, which is within the Virtual World Second Life [4]. The laboratory simulates the facilities, equipment, documentation and tools afforded by a real laboratory of the pharmaceutical industry that complies with GLP (Good Laboratory Practices) regulations and has been used in teaching since 2011 to train graduate students in subjects related to Quality in the Pharmaceutical Industry [5, 6]. The reason for building a virtual scenario of this type of laboratory is the economic cost and physical space that must be used, non-asumible for a university or any other institution that would not get any direct economic profit from it.

The training that students receive in this virtual scenario is based on learning methodologies, standards and audit processes in laboratories of the pharmaceutical industry in a practical way (and not merely theoretically, as had been usual), through interaction with a 3D scenario that replicates a laboratory that represents a real case. To do this, students act within virtual facilities as true external auditors who are responsible for assessing compliance with the laboratory of these GLP measures, generating a report outlining compliance (or not) of the rules at the end of practical sessions, flaws identified in it (there are always a number of deficiencies brought intentionally by teachers), critical nature, etc. This type of training has been included since 2011 as part of the subject Quality Assurance in the Laboratory Analysis in the Pharmaceutical Industry within the Master in Drugs Evaluation and Development of the University of Salamanca (Spain).

III. The Problem: Proposal, Objectives and Requirements

This section presents the problem proposed by teachers and technical personnel responsible for the Usalpharma Lab and the specific objectives to be met by the solution, as well as the software requirements developed to solve the problem.

A. The proposal of the problem

Between 2011 and 2013 this scenario was used in teaching under the direct supervision of teachers of the course and staff involved in the development of 3D environments. That is, practice work used to be developed in the laboratory (in a single session and call for all students), was guided by a teacher and was supervised by the rest of the team, so that if a student had any questions or appreciation of the 3D scenario, content resources, possible variations or disagreements regarding present rules or any other aspect they could immediately make a query, getting instant feedback and sometimes very specific help. This stage, as far as results are concerned, was very positive for the initiative, students got good grades in their practice and showed satisfaction with the methodology and resources used in this virtual learning [7].

After this first stage in the implementation of this type of system to support teaching, the teaching innovation team wanted to go a step further, through implementing a system that would *replace* the teacher's guide during the students' practical sessions, controlling whether students were doing or not the practice within the virtual world, as well as to include a number of requirements on the feedback that students receive when participating in their practical work. That is, a system that allows taking advantage of the potential of virtual worlds in terms of the ability of autonomous learning, the possibility of not depending on the time and physical space (time, physical location of students) to leverage resources provided by virtual learning environments as well as presenting better immersive experiences to students within the 3D environment [8-12].

B. Specific objectives

Then, the proposed objectives for this system (both objectives proposed in the first version of the system, as well as those proposed for successive versions) are as follows:

- The system should be able to monitorize what students do in the virtual laboratory, discriminating who clicks on any object, when they do, and what object or part of the particular installation is used.
- This system should allow the teacher to know the actions carried out by each student, showing him or her some kind of report associated with each student or user registered by the system. These actions can be consulted unformatted (or in minimum formats such as lists or tables) and be filtered using temporal parameters (date range), by object type, etc.
- The system should control if the student is carrying out the necessary checks on those sensitive elements to breach the rules, especially those that teachers marked incorrectly on the virtual stage. The result of these checks should be included in the report on the activity of each student.
- The system should provide general reports on the activity of a group of users at any given time, not only individually for each registered user.

- The system must be able to transmit real-time feedback to students on their progress in their practical work.

C. *Software requirements*

Among the determined requirements for this system the following were mainly specified:

- The system should store for each action performed in the virtual world (click on objects) the name of the user performing the action, the name or descriptor of the object that receives the action, the specific time when that has been done (timestamp) as well as the specific location where the action was performed.
- The system should show teachers a report of students' actions regardless of device and operating system from which the report was accessed.
- The system should provide in-world feedback to students, that is, it must have the necessary tools not only to extract information from Virtual World, but also to provide input.
- The system must efficiently store the required data. The data layer should allow varying the model and the data types that support, in anticipation of possible changes suffered by the system in the future, or additional features that should be contemplated.
- The system must be prepared to incorporate new sources of data (student grades, personal data, integration with other platforms, etc.) in the possible future.

Regarding other possible requirements to be specified (functional, non-functional or others for information) as the technology that must be employed, modeling specific data, or aspects of extensibility and modularity of the system, all the team members agreed to request these aspects to the engineers,, so they evaluated what might be the best option in each case.

IV. Solution and Product Developed

In this section are detailed three fundamental parts of the search process of a proper solution to the proposed problem: previous considerations, the theoretical solution proposed and the final developed product that is currently on production.

A. *Previous considerations to the design of the solution*

Before designing and developing a viable and optimal solution for this problem, researchers had to take into account a number of specific preliminary considerations of the problem to be solved, among which are highlighted:

- The problems posed by the interconnection of a private and closed system such as Second Life with any external platform. Not only does it present problems regarding integration, there are no public resources, logs, APIs or other tools to integrate with other systems [13, 14] beyond basic information on number of users per day, etc. [15]. It presents additional problems because of the rigid characteristics regarding the ownership of the 3D objects, which hinders the usage for the described purpose. Even more, Second Life is provided "as is", which prevents any claims about failures, changes, etc. and requires users to modify the system behaviour [16].
- The system that would be proposed as a solution must be able to coordinate the response to the whole problem, from data collection to the presentation of the same on any device or within the Virtual World. This implied drawing on well-known open technologies, standards that have an almost complete presence in every possible system. These standards must be present in the technological solution as well as communication protocols, data collection, etc. [17, 18].
- The technology that supports the management of processes, such as layer data management, must be able to withstand peaks of activity (many students interacting simultaneously in the 3D laboratory), possible massive load at certain time intervals (while carrying out the practical work) and concurrency in the interaction with the resources that provides performance without excessive erosion.

B. *Theoretical solution proposed*

Since 2013 the authors have been working to address the problem in a satisfactory manner, and in early 2014 a viable solution was finally proposed [14] (at least from a theoretical point of view) that provided an adequate framework and a proof of concept for the final development of a product. This solution could remedy the problem regarding integration of a system that monitorizes, analyzes, reports and helps the evaluation of teaching activities within the virtual laboratory Usalpharma Lab.

This solution is based on an architecture that could be deployed in a cloud environment so that the architecture, based on the typical client-server schemes, consisted of a number of layers (data collection, data persistence of analysis, presentation, etc.) connected to each other in the same way (client-server depending on which requires the services of other interactions), enabling the deployment of each layer even in different clouds (with the potential to scale only those resources required, apply different technologies to each layer, etc.) due to the fact that the layers simply interact among them using services (Figure 1).

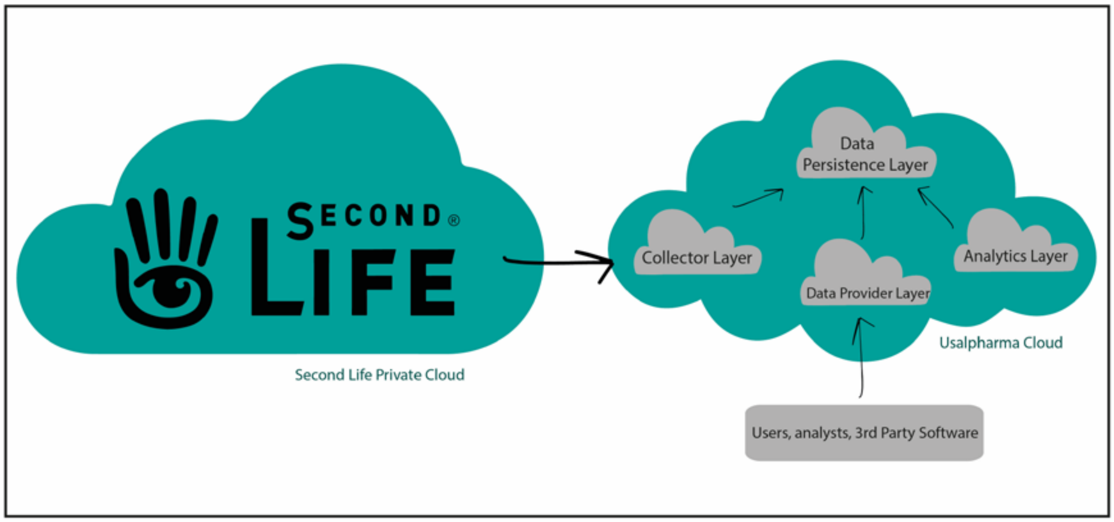


Fig. 1. Simplified representation of the architecture of the proposed solution [14]

Regarding the standards and technologies proposed in this theoretical solution, the following can be outlined:

- Use of a (simplified) variation of the semantic standard RDF and its triples [19, 20] to transmit structured and rich information about the interaction in the Virtual World to the data collection layer.
- Use of the HTTP protocol [21] for the transmission of messages between the Virtual World and the layers of data collection and provision (in both directions, the Virtual World to the system, and the system to the Virtual World), specifically using the part of the protocol concerning operations, requests, status codes, and headers.
- Intensive use of web technologies to return information to users (teachers and students) regardless of device, operating system or software (such as Second Life viewer).
- Using NoSQL documents-oriented databases (MongoDB [22] specifically) that allow modification of data architecture transparently and without penalty arising from the persistence of the layer architecture (only possible penalties in the logical layer derived from changes in its own data models [23]), and maintaining good performance in information processing, and high scalability in production environments. [24].
- Use of server technologies that have demonstrated their capacity for data processing HTTP requests and adequate performance under heavy demand. In this case, the technology proposed was the web framework Django [25], which develops robust solutions using Python and has many libraries of its own (and others of the language) to facilitate the development of stable agile projects.

This solution provides a flow of activity as proposed in the activity diagram shown in Figure 2, which is an updated version of the software's architecture of the version currently implemented. In this diagram, the interaction can be seen between the system and the actors (stakeholders) involved in it, and serves as a summary of the workflows of the whole process.

C. *Product developed*

In this third part of the section are detailed the various major components of the product developed from the theoretical proposal for solving the problem. This product consists of three main parts, architecture and server itself, as well as Web clients (for teachers) and client embedded in Second Life (for students):

Software Architecture: As previously discussed, it is implemented using the web framework Django, and consists of a series of applications that are responsible for collecting the HTTP requests that launches the Second Life client for recording data interaction, cleaning and processing such information, contacting the persistence layer data and database MongoDB data, serving requests such as "display information" depending on the context / client from which they are requested (the same information is not served when the user launches requests from the web application or from the Virtual World), calculating data associated with user interaction: measures of the time spent by each student to complete their practical work [26], calculating the most relevant key points achieved [27], reporting for each user or group of users, etc. These applications and the multiple layers composed architecture are responsible for maintaining the logic of the whole system as well as the bulk of the functionality.

Web Client: From the Web client, teachers can consult all data concerning students and their work progress. Among the metrics that can be known should be available the number of collected interactions, time spent by students performing, the achievement by the students of the inspection of the various elements (review of safety measures laboratory, review of equipment to be audited, documentation, etc.) as well as comprehensive reports for each class of students, so that a teacher can know in a detailed way the performance of each student, or acquire an overview of the group of students in their practicals. This web client is developed to be used from any device (PC, tablets, smartphones, etc.). Figure 3 presents some data visualization of various metrics that teachers can know about student performance in Usalpharma Lab.

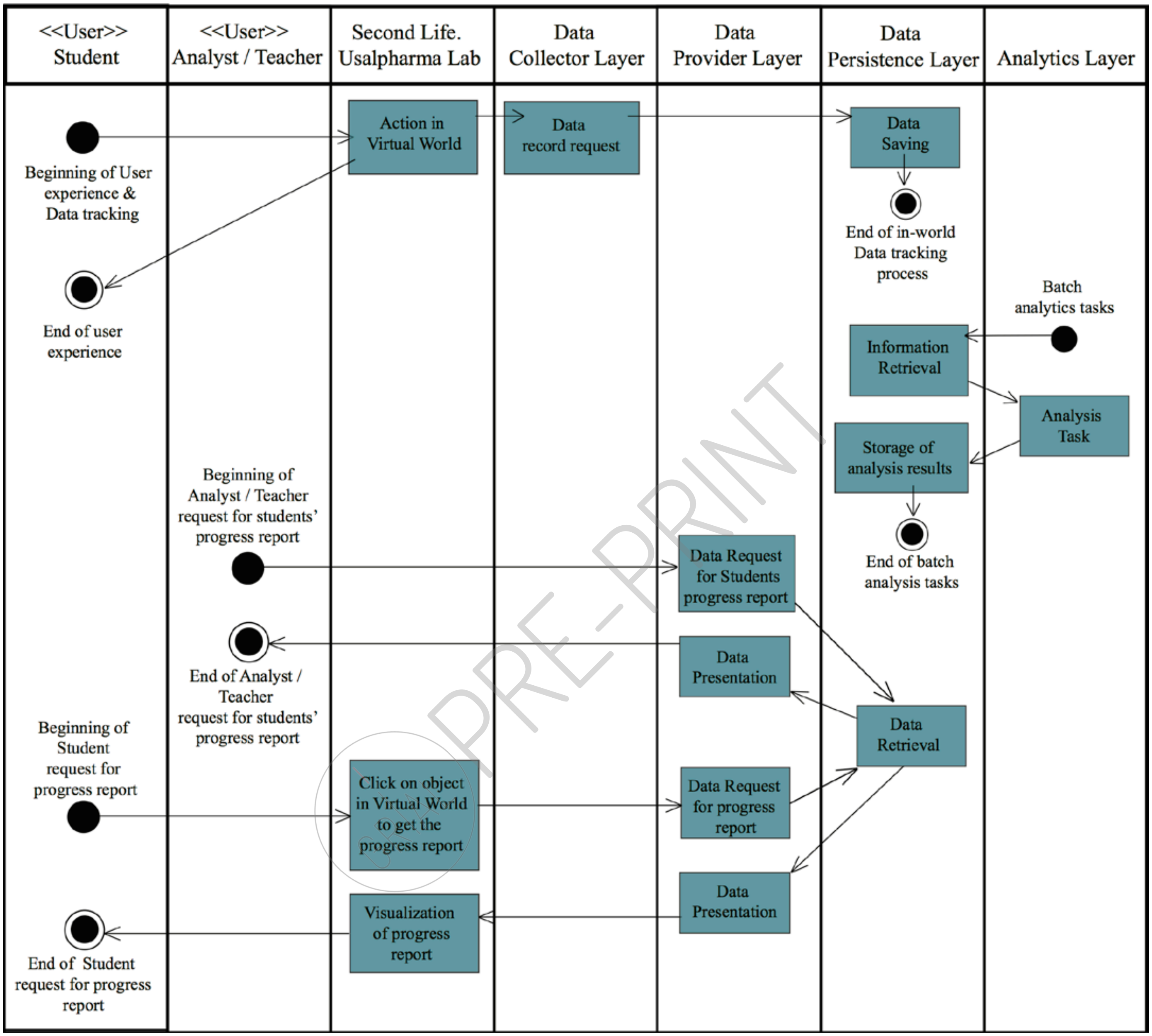


Fig. 2. Activity diagram among users, Virtual World and system. Updated version of the proposed in [14]

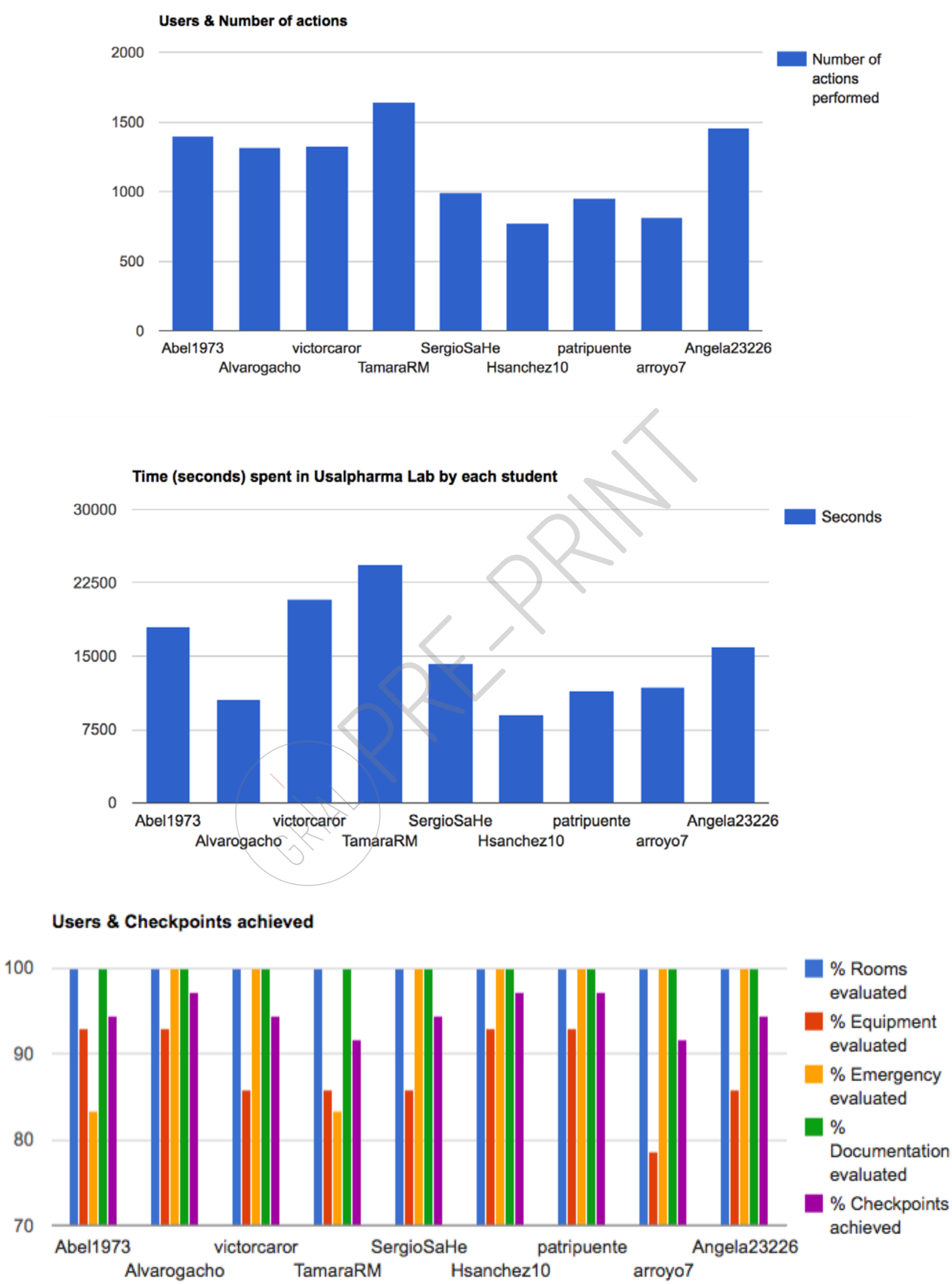


Fig. 3. Some data visualization presented in the Web client. They show different metrics related to student activity in the 2015 practice

Second Life client: This client is somewhat different from the web client because it can be used by any avatar (user in Virtual World) that is in the laboratory. The client is used through interaction (click) on an object in the laboratory, so that the student will see a dialog showing in real time the percentage of critical points assessed, both in the day when a query is made and, in general, making suggestions on what the student should re-audit. In the first year of practice the system showed only feedback about the particular day of the request, without giving an overview of the development of the practical on different days given to audit. Figure 4 shows this type of activity report.

At the same time, the Second Life client has an even more essential role, and that is to send data to the system, so that it registers every evidence of interaction between the user and the virtual laboratory (data entry to the platform). To do this, all 3D objects to be audited by students in the virtual laboratory have an associated script that throws a HTTP request to the application, which is responsible for collecting information about user interaction to record the action that is occurring at the moment. This request is carried out silently and transparently for the user; the object itself is responsible for collecting the data on the date and the exact moment of interaction (timestamp), with what the user has interacted, where in the virtual lab it is, etc., so that the user is not aware of being transmitting that information to the system, although all students are previously notified of the monitoring carried out in the laboratory.

D. *Evolution and improvement of the proposal*

As mentioned, in 2014 a first formal proposal was made and published, explaining how to address the problem [14]. In that first proposal a system was presented that would allow the collection of learning evidences and that could offer teachers some level of information about what students performed within the virtual laboratory environment. It also focused on the problems that such proposal must face from the point of view of software engineering to obtain information from the virtual world and take proper advantage of it as an evidence of learning by defining and using RDF simplified schemes, making data collection through structures and information architectures flexible enough to vary the type of information retrieved, and allowing the interoperability of the data schema to other systems (new developments or third-party systems) in a simple and transparent way.

After this initial proposal, the project came under development and, based on the first results and observed possibilities, the authors saw the opportunity to expand the system functionality, including a few months after the first enforcement and feedback mechanisms within the 3D laboratory [27] and opening the way for learning personalization and adaptivity of learning (referred as part of future work, as will be discussed at the end of the paper).

Once implemented these features and having a sufficiently functional system, it was tested with real users (years 2013-2014 and 2014-2015). These tests were helpful to verify the actual usefulness of the system and its operation. Likewise, once the system had been put to production, the authors made improvements in the reporting of users (students in the Virtual World, teachers using the Web client, etc.).

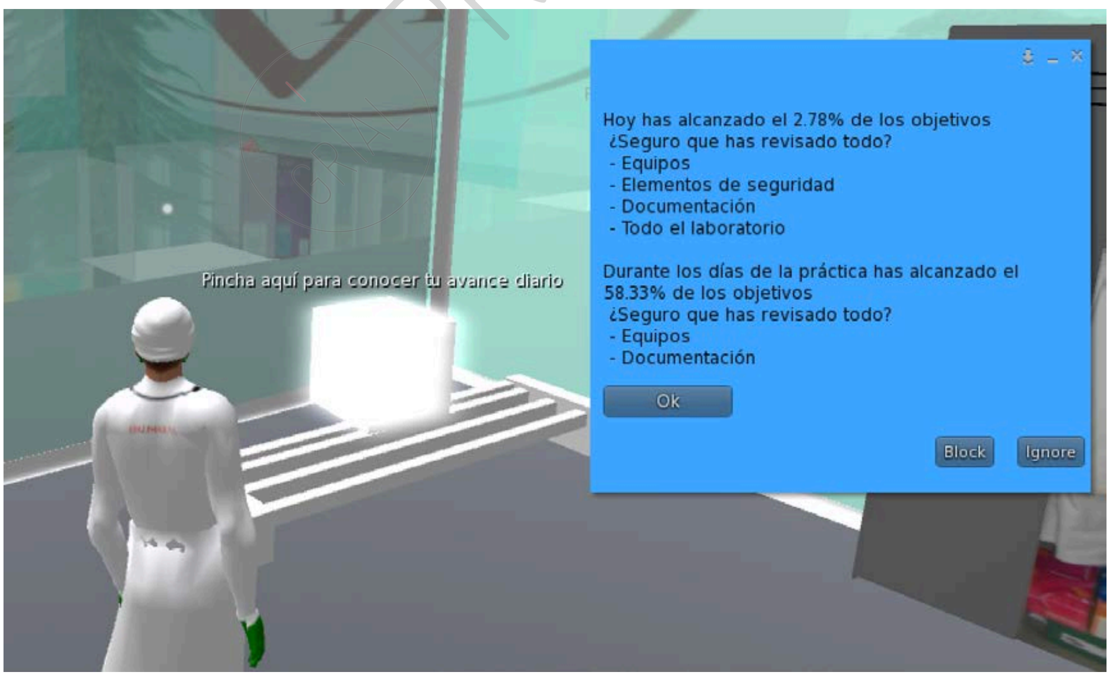


Fig. 4. Real-time information that students can get in Second Life about the audit performed. In the Spanish text that appears in the figure, the system informs the student about the performance (% of the goals in the day and in all the days of the practice) and suggests the different elements where the student has not completed the audit.

V. *Exploitation and Results*

The solution developed has been in production since February 2014, monitoring continuously since that time (excluding time spent on upgrading the system to the latest versions) any relevant activity that has occurred in the laboratory. Specifically, the system has already monitorized two real practicals with students, one in the course 2013-2014 (between 6 and 24 March 2014) and another in 2014-2015 (between 5 and 12 March 2015) of the aforementioned subject Quality Assurance in the Laboratory Analysis in the Pharmaceutical Industry within the Master in Drugs Evaluation and Development of the University of Salamanca (Spain). The system has also monitorized the interaction that other users have had during the time elapsed between such practicals. These have involved 18 students (9 in the 2013-2014 year, 9 in the 2014-2015 year) that were enrolled in the subject.

In such practice sessions, these tools presented here have been taken into account for the assessment of the subject, because teachers have decided to give a part of the evaluation of practice to the data provided by the system. Specifically, 25 percent of the practice mark directly depends on the percentage of critical points assessed by the student in the laboratory (% checkpoints achieved in Figure 3). The other 75% of the practice mark corresponds to the audit report presented by students at the end of their practice sessions, and where they really show their knowledge about the GLP regulations and the elements of the laboratory that meet them or not. Other data taken into account when assessing the practicals are: a) if the student has carried out work on the agreed dates, b) if the student has made more than 1 hour of audit (in one or more sessions), c) if they have reached more than 12% of checkpoints revised. The teaching staff sets these three elements as the minimum requirements to show that students have actually audited and evaluated the laboratory.

Among the different results that can be shown to illustrate the operation of the entire system, it is possible to distinguish between two vectors of data to help test whether the system works, and to indicate if the system may be accepted for its actual use in teaching:

- Data concerning the validity of the system to help students practice and somehow replace a personal support teacher in the virtual lab.
- Data about measurements that show the usefulness of a system of collection, sorting and processing of data in a case like this.

In the first case (*validity of the system to help students practice and somehow replace personal support teachers in the virtual lab*), in order to judge the results it is possible to compare the average marks obtained by students in the different years that the virtual laboratory has been used for practicals (years with direct support of the teachers as well as years with the implementation of this architecture and software tools). Table 1 shows a summary of the marks:

TABLE I

STUDENTS AND THE AVERAGE MARKS CORRESPONDING TO THE COURSES WHEN PRACTICALS WERE REALIZED

Course	Number of students	Average mark obtained (0 to 10)	Standard deviation
2011-2012	14	7.28	1.38
2012-2013	16	7.5	1.27
2013-2014	9	7.81	0.7
2014-2015	9	8.54	0.56

Based on these results, it is possible to see how students have increased their average mark (and decreasing the standard deviation) in each year, showing no significant changes in the trend between the different years –namely, the disparities observed between those years where the practice was conducted as a single session with teachers' guidance and the other years where the practice was aided using the complementary technological tools developed within the scope of this project. As will be discussed in the discussion section, these data do not have to convey that the use of this architecture and the tools provided improve the results of practical sessions (this may be due to many factors, such as the decline in number of students, the possible prior habituation of students to similar tools or 3D contexts, etc.), but it illustrates how a trend of improving results is maintained even though the method of help, support and evaluation of the practicals varies.

As for the second information vector, about the usefulness of this type of system for monitoring, reporting results and aiding in the evaluation and development of practical work, a series of metrics can be displayed collected by the platform that illustrate the reader about the volume of information that the system has collected in these two academic years (2013-2014 and 2014-2015, specificly between 18 February 2014 and 20 March 2015), saving teachers from:

- 24276 interaction evidences (clicks in the laboratory).
- 249 sessions by 47 different users (teachers, students and visitors to the facilities).
- 354269 seconds of interaction registered (total time of use of the system by the users, measured in seconds). It is more than 98 hours of interactions retrieved, stored and analyzed for use by teachers. Of this usage time, 137454 seconds (over 38 hours) of use of the laboratory correspond with students who carried out their practical work in the year 2013-2014 and 136581 seconds (almost 38 hours, again) of usage correspond to the time students who participated in the year 2014-2015 (in total, these students have registered almost 78% of total usage time in the laboratory).

These data are only a sample of what the Usalpharma architecture has contributed to the development of practice and the enormous task of knowing what happens in the virtual laboratory These data are particularly relevant when analyzed in context: the data retrieved represent all the users' interactions over a long period of time, regardless of the time zone of the visitor, the specific time of connection. The data provide metrics that teachers would never know or control without the help of the system.

VI. Discussion

This section focuses on commenting the main points shown in the article, highlighting both the lights and the shadows cast by this experience, the tools developed, and the utility that the system acquires in a learning context such as the one shown. To do so, a number of subsections are set, each of which presents a question, key in the author's view, which will be answered as concisely as possible:

A. *Is this proposal innovative?*

It is true that in the field of serious games there are already initiatives for the interconnection with systems like MOOCs, PLEs, VLEs, Learning Analytics platforms, etc. [28-33], but in the realm of virtual worlds there are not many such initiatives [34], since a review of the current literature reveals that there is a tendency in the approach to proposals for the integration of different systems like VLEs, PLEs, architectures and applications together with Virtual Worlds [28, 35-39], but not with systems that analyze the performance of students within virtual practicals, or analyzing (however basic it may even be considered) the interaction and the evidences collected from student usage of 3D learning scenarios and practices. The literature mainly describes the interconnection of Virtual Worlds environments like OpenSim or Second Life with widespread platforms in the current teaching (Moodle, mainly) by using Sloodle and other platforms, so that these 3D environments extend their horizons of educational usage based on the support of systems and platforms where teachers already perform constant work creating materials and using them as an aid to teaching.

In fact, the authors of this paper believe that this system is a true innovation in the field of teaching in virtual worlds (and even more in the field of education in Health Sciences) because of its features and vision outside the mere connection with systems that provide learning objects or materials created by teachers, but as an aid to the evaluation and understanding of how students perform in such complex and dynamic environments like these [40]. Even in the most recent literature it is possible to find authors (Griebel *et al* [41], for example) that describe the proposal of this Usalpharma architecture as an interesting contribution that can open up new possibilities in the use of external systems completing the educational use of virtual worlds through the evidence collection system, the real-time responses both within the Virtual World and outside it, and its possibilities of use for the evaluation of activities. Therefore, for the authors it is possible to consider it a real, innovative and viable proposition for help in teaching. The interconnection of this architecture (using the services already available and others to be built) with other applications extends its field of education and eLearning, and could help to achieve better results in learning scenarios arranged in 3D environments of this type.

B. *Is this proposal helpful?*

The usefulness of such systems, from the point of view of the authors, depends primarily on the problem that it helps to solve. In the case of Usalpharma laboratory, this proposal provides information that cannot be achieved otherwise within a virtual world like Second Life, but it is necessary to consider whether this information is vital or not in a process of teaching and evaluation of students. In this case, the authors agree that, in the present case, the information is not vital, since evaluation depends mostly on the report submitted by the students (it represents 75% of the final mark); therefore, this system only provides extra assessment elements. What must be noted is that, in this particular case, those extra elements are also designed to establish a right to assessment or not (the reader will remember the conditions set for the assessment of the practical, of at least 12% of proven elements and a minimum time of interaction with the virtual laboratory of one hour). Moreover, the authors also wish to state that the contribution made by the system in the form of information for students and teachers is a faithful overview of what happens in a environment like a virtual world, allowing teachers to measure aspects such as dedication (hours used), effort (number of sessions, number of actions to detect all elements that must be evaluated) or persistence (number of sessions used, evolution of the completion of the activity in each session, etc.), for example. This type of data, although not taken into account in the specific assessment for this practical work (and therefore the architecture is not decisive in the present evaluation) could be considered as catalyst values for the final evaluation. The data were part of the rubric for assessing practice work, giving specific metrics to aspects that are not typically measured in a clear way. In summary, although the system has not a vital utility for the practical work, it is increasingly being consolidated as a helper to the teacher. Additionally, the system is aligned with the current trends of tools, applications and utilities that are opening new paths in new learning contexts that will be relevant in technology-enhanced learning in the future..

C. *Does it help to improve teaching?*

In this sense, it is difficult to provide a clear answer. By the data in Table 1, it is possible to say that the shift of paradigm and support in this case from classroom guidance to a system that provides feedback to students automatically, has not broken the trend of improved results, even when the results were improved significantly in the last year. The paradigm shift provided by this system is helping to improve teaching in a context such as a Virtual World, allowing students to experience completely immersive environment like Second Life and avoiding the guidance of the teacher in real time. However, venturing this conclusion would be going too far as it is possible that other factors affect results. Among the factors that could be relevant for the improvement of the results, can be highlighted the smaller number of participant students in the last years, the students' better predisposition towards such tools, the previous experience of some students with this kind of audit simulation, etc.

In any case, time will determine whether this system or others like this will significantly and unequivocally improve the students' learning and results, or automation of certain tasks previously performed by the teacher, are not really replaceable, and these systems simply must serve teachers as an aid in the acquisition of knowledge and decision making.

D. *Can the system have more applications in addition to those described?*

This proposal is, even today, a proof of concept on the use and application of such systems and platforms related to eLearning. The authors strongly believe that this proposal could have a more intensive usage and increased functionality through integration with other systems and eLearning platforms, besides using more advanced algorithms and analysis techniques than those currently applied, so that they could get to perform complex learning analysis, behavioral analysis (which is already done in Virtual Worlds in other ways [40]), determination of the learning path for each student, personalizing learning, development of adaptive systems within Virtual Worlds, copy-detection among practices performed in 3D environments, etc.

VII. CONCLUSIONS AND FUTURE WORK

In this last section are presented the conclusions and future lines of work that arise within the project Usalpharma.

A. *Conclusions*

This article shows how a software architecture has been conceived, designed, implemented and tested that helps support an educational activity in a learning environment like a Virtual World. The article delves into the problems that such systems can solve as well as the benefits, profit, exploitation thereof, or the improvement achieved from current systems and methodologies.

From a standpoint of software engineering, this paper describes how a team of engineers has been able to develop an innovative solution to a problem of some complexity, showing what tools and methods they followed to overcome the difficulties, how they planned the development and use of the solution from a multi-platform view.

From a standpoint of eLearning, this work shows how a multidisciplinary team has been able to design, build and put in use a solution to a real and specific learning problem of graduate students from the branch of knowledge of Health Sciences and Pharmacy. In addition, from this point of view, this work shows the results that have been achieved in the first two years that the proposal has been used, and how these results have been satisfactory and provide real value to the activity of teachers and students.

B. *Future work*

As future work to improve and expand this architecture and its use, it is possible to distinguish three major ways:

- In the short-term the authors plan to increase the metrics shown and calculated (e.g., see how many times a student consults his or her progress, the distribution over time of the students' effort with their practicals, etc.). The authors will also continue using the solution described in this article in successive editions of the subject Quality Assurance in the Laboratory Analysis in the Pharmaceutical Industry within the Master in Drugs Evaluation and Development of the University of Salamanca and expand its use to other subjects that use this type of teaching methodologies through Virtual Worlds, to continue testing the solution in different contexts and assess more users (getting better measures on usefulness of these tools for teachers).
- In the medium term, the authors want to create a set of APIs and utilities, allowing the interconnection of this system with other platforms, so that the proposed architecture will serve as a facilitator of data and metrics that can be expanded, used or integrated into current or future eLearning systems to help close the loop between the various technological tools that help learning any subject.
- In the long term, the authors plan the possibility of integrating the system with other standard RDF specifications and recommendations. Among them they propose the integration of the description of the interactions, the description of objects and users through Open Linked Data [42], following the schemes of the Linked Data Platform recommendation of the W3C [43]. Thus, the architecture and its components could eventually contain all the information concerning the activity and can have descriptions associated with each object, each target, maintaining user profiles under specific linkable URIs, etc., allowing mapping the domain of the problem and exploiting all the information from clients and developed platforms as well as from potential third-party applications connected to services that will provide the architecture.

ACKNOWLEDGMENTS

The authors wish to thank the Vice Chancellor for Academic Policy at the University of Salamanca (Spain) for funding the Teaching Innovation projects ID2013/201 and ID2014/0164 that supported the development of the software architecture and software solutions presented in this paper. Also, the author Juan Cruz-Benito would like to thank the European Social Fund and the Consejería de Educación of the Junta de Castilla y León (Spain) for funding his pre-doctoral fellow contract.

REFERENCES

[1] F. J. García-Peñalvo, M. Á. Conde González, and S. Bravo Martín. (2009, 10/02/2015). *Ingeniería del Software: Introducción*. Available: http://ocw.usal.es/ensenanzas-tecnicas/ingenieria-del-software/contenidos/Tema0-Sumario-1pp.pdf

[2] IEEE, "IEEE Standard Glossary of Software Engineering Terminology," in *IEEE Std 610.12-1990*, ed, 1990, pp. 1-84.

[3] J. Cruz-Benito, R. Therón, and F. J. García-Peñalvo, "Software Architectures Supporting Human-Computer Interaction Analysis: A Literature Review," in *Learning and Collaboration Technologies*, ed: Springer International Publishing Switzerland, In press.

[4] Linden Lab. (2014, Feb 28, 2014). *Second Life*. Available: http://secondlife.com

[5] C. Maderuelo, A. Martin-Suarez, J. S. Pérez-Blanco, H. Zazo, J. Cruz-Benito, and A. Domínguez-Gil, "Facility-based inspection training in a virtual 3D laboratory," *Accreditation and Quality Assurance,* vol. 19, pp. 403-409, 2014/10/01 2014.

[6] C. Maderuelo, J. S. Pérez-Blanco, H. Zazo, J. M. Armenteros del Olmo, J. Cruz Benito, and A. Martín-Suárez, "Auditing Training Practice To Postgraduate Pharmacy Students In A Virtual World," presented at the 19th Annual Conference of the European Association of Faculties of Pharmacy (EAFP 2013), Ankara, Turkey, 2013.

[7] C. Maderuelo, J. S. Pérez-Blanco, J. M. Armenteros del Olmo, H. Zazo, D. Sánchez Vicente, J. Cruz-Benito*, et al.*, "Satisfacción de estudiantes de Postgrado con una práctica en Second Life," presented at the Congreso Virtual Mundial de e-Learning 2012, 2012.

[8] L. Jarmon, K. Lim, and B. Carpenter, "Introduction: Pedagogy, education and innovation in virtual worlds," *Journal of Virtual Worlds Research,* vol. 2, pp. 3-4, 2009.

[9] J. Shen and L. B. Eder, "Intentions to use virtual worlds for education," *Journal of Information Systems Education,* vol. 20, p. 225, 2009.

[10] M. Varvello, S. Ferrari, E. Biersack, and C. Diot, "Exploring second life," *IEEE/ACM Transactions on Networking (TON),* vol. 19, pp. 80-91, 2011.

[11] M. Virvou and G. Katsionis, "On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE," *Computers & Education,* vol. 50, pp. 154-178, 2008.

[12] Q. Wang, "Design and evaluation of a collaborative learning environment," *Computers & Education,* vol. 53, pp. 1138-1146, 12// 2009.

[13] J. Cruz, R. Therón, E. Pizarro, and F. J. García-Peñalvo, "Knowledge Discovery in Virtual Worlds Usage Data: approaching Web Mining concepts to 3D Virtual Environments," *Proceedings Fourth International Workshop on Knowledge Discovery, Knowledge Management and Decision Support (Eureka-2013),* 2013.

[14] F. J. García-Peñalvo, J. Cruz-Benito, C. Maderuelo, J. S. Pérez-Blanco, and A. Martín-Suárez, "Usalpharma: A Cloud-Based Architecture to Support Quality Assurance Training Processes in Health Area Using Virtual Worlds," *The Scientific World Journal,* vol. 2014, 2014.

[15] Linden Lab. (2015, 28/07/2015). *APIs and Web Services Portal*. Available: http://wiki.secondlife.com/wiki/APIs_and_Web_Services_Portal

[16] Linden Lab. (2015, 15/03/2015). *Second Life. Terms of Service*. Available: http://www.lindenlab.com/tos

[17] Á. del Blanco Aguado, J. Torrente, I. Martínez-Ortiz, and B. Fernández-Manjón, "Análisis del Uso del Estándar SCORM para la Integración de Juegos Educativos," *Revista Iberoamericana de Tecnologías del/da Aprendizaje/Aprendizagem,* p. 118, 2011.

[18] A. del Blanco, A. Serrano, M. Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, "E-Learning standards and learning analytics. Can data collection be improved by using standard data models?," in *Global Engineering Education Conference (EDUCON), 2013 IEEE*, 2013, pp. 1255-1261.

[19] World Wide Web Consortium (W3C). (2012, 15/03/2015). *RDF Interfaces. A set of basic primitives and a low-level API for working with RDF data*. Available: http://www.w3.org/TR/2012/NOTE-rdf-interfaces-20120705/

[20] World Wide Web Consortium (W3C). (2004, 15/03/2015). *RDF Semantics. W3C Recommendation 10 February 2004*. Available: http://www.w3.org/TR/2004/REC-rdf-mt-20040210/

[21] World Wide Web Consortium (W3C). (2014, 15/03/2015). *HTTP - Hypertext Transfer Protocol*. Available: http://www.w3.org/Protocols/

[22] MongoDB. (2015, 15/03/2015). *Website*. Available: https://www.mongodb.org/

[23] J. Cruz-Benito, F. J. García-Peñalvo, and R. Therón, "Defining generic data collectors for Learning Analytics: Facing up the heterogeneous data from heterogeneous environments," presented at the International Conference on Advanced Learning Technologies (ICALT) 2014, Athens, Greece. , 2014.

[24] MongoDB, "Top 5 Considerations When Evaluating NoSQL Databases," ed, 2013.

[25] Django Software Foundation. (2015, 15/03/2015). *Django Web Framework*. Available: https://www.djangoproject.com/

[26] J. Cruz-Benito, F. J. Garcia-Peñalvo, R. Theron, C. Maderuelo, J. S. Perez-Blanco, H. Zazo*, et al.*, "Using software architectures to retrieve interaction information in eLearning environments," in *Computers in Education (SIIE), 2014 International Symposium on*, 2014, pp. 117-120.

[27] J. Cruz-Benito, R. Therón, F. J. García Peñalvo, C. Maderuelo, J. S. Pérez-Blanco, H. Zazo*, et al.*, "Monitoring and feedback of Learning Processes in Virtual Worlds through analytics architectures: A real case," in *Sistemas y Tecnologías de Información. Actas de la 9ª Conferencia Ibérica de Sistemas y Tecnologías de Información*. vol. I Artículos, Á. Rocha, D. Fonseca, E. Redondo, L. P. Reis, and M. P. Cota, Eds., ed Barcelona, España, June, 18-21, 2014: AISTI (Asociación Ibérica de Sistemas y Tecnologías de Información), 2014, pp. 1126-1131.

[28] L. Stoicu-Tivadar, V. Stoicu-Tivadar, D. Berian, S. Drăgan, A. Serban, and C. Serban, "eduCRATE-A Virtual Hospital Architecture," *Studies in health technology and informatics,* vol. 205, pp. 803-807, 2013.

[29] Á. Serrano-Laguna, J. Torrente, P. Moreno-Ger, and B. Fernández-Manjón, "Application of learning analytics in educational videogames," *Entertainment Computing,* vol. 5, pp. 313-322, 2014.

[30] Á. Serrano-Laguna, J. Torrentea, B. Maneroa, Á. del Blancoa, B. Borro-Escribanoa, I. Martínez-Ortiza*, et al.*, "Learning Analytics and Educational Games: Lessons Learned from Practical Experience," in *Games and Learning Alliance*, ed: Springer, 2014, pp. 16-28.

[31] J. Cruz-Benito, O. Borrás-Gené, F. J. García-Peñalvo, Á. Fidalgo Blanco, and R. Therón, "Detection of Non-Formal and Informal Learning in learning communities supported by social networks in the context of a Cooperative MOOC," in *Proceedings of the XVII International Symposium on Computers in Education (SIIE'15) (Setúbal, Portugal. 25-27 November 2015)*, M. d. R. r. Rodrigues, M. n. Llamas Nistal, and M. Figueiredo, Eds., ed Portugal: IEEE, 2015.

[32] J. Cruz-Benito, O. Borrás-Gené, F. J. García-Peñalvo, Á. Fidalgo Blanco, and R. Therón, "Extending MOOC ecosystems using web services and software architectures," in *Proceedings of the XVI International Conference on Human Computer Interaction*, ed Vilanova i la Geltrú, Spain: ACM, 2015, pp. 1-7.

[33] F. J. García-Peñalvo, J. Cruz-Benito, O. Borrás-Gené, and Á. Fidalgo Blanco, "Evolution of the Conversation and Knowledge Acquisition in Social Networks Related to a MOOC Course," in *Learning and Collaboration Technologies*. vol. 9192, P. Zaphiris and A. Ioannou, Eds., ed: Springer International Publishing, 2015, pp. 470-481.

[34] J. Cruz-Benito, C. Maderuelo, F. J. García-Peñalvo, R. Therón, J. S. Pérez-Blanco, H. Zazo*, et al.*, "Usalpharma: Una arquitectura software al servicio del aprendizaje en Mundos Virtuales," *VAEP RITA,* vol. 3, pp. 148-159, 2015.

[35] D. Livingstone and J. Kemp, "Integrating web-based and 3D learning environments: Second Life meets Moodle," *Next Generation Technology-Enhanced Learning,* vol. 8, 2008.

[36] J. W. Kemp, D. Livingstone, and P. R. Bloomfield, "SLOODLE: Connecting VLE tools with emergent teaching practice in Second Life," *British Journal of Educational Technology,* vol. 40, pp. 551-555, 2009.

[37] E. P. Lucas, J. Cruz–Benito, and O. G. Gonzalo, "USALSIM: learning, professional practices and employability in a 3D virtual world," *International Journal of Technology Enhanced Learning,* vol. 5, pp. 307-321, 2013.

[38] A. Konstantinidis, T. Tsiatsos, S. Demetriadis, and A. Pomportsis, "Collaborative learning in OpenSim by utilizing sloodle," in *Telecommunications (AICT), 2010 Sixth Advanced International Conference on*, 2010, pp. 90-95.

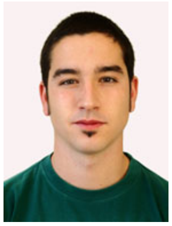
[39] P. Kallonis and D. G. Sampson, "A 3D Virtual Classroom Simulation for Supporting School Teachers Training Based on Synectics-" Making the Strange Familiar"," in *Advanced Learning Technologies (ICALT), 2011 11th IEEE International Conference on*, 2011, pp. 4-6.

[40] J. Cruz-Benito, R. Therón, F. J. García-Peñalvo, and E. Pizarro Lucas, "Discovering usage behaviors and engagement in an Educational Virtual World," *Computers in Human Behavior,* vol. 47, pp. 18-25, 6// 2015.

[41] L. Griebel, H.-U. Prokosch, F. Köpcke, D. Toddenroth, J. Christoph, I. Leb*, et al.*, "A scoping review of cloud computing in healthcare," *BMC medical informatics and decision making,* vol. 15, p. 17, 2015.

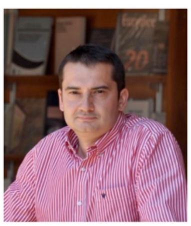
[42] C. Bizer, T. Heath, and T. Berners-Lee, "Linked Data - The Story So Far," *Int. J. Semantic Web Inf. Syst.,* vol. 5, pp. 1-22, / 2009.

[43] World Wide Web Consortium (W3C). (2015, 15/03/2015). *Linked Data Platform 1.0. W3C Recommendation 26 February 2015*. Available: http://www.w3.org/TR/2015/REC-ldp-20150226/



Juan Cruz-Benito received the M.Sc. degree in intelligent systems from the University of Salamanca, Spain, in 2013, where he is currently pursuing the Ph.D. degree in computer sciences. He is one of the youngest members of the Research Group Interaction and eLearning, where he specializes in software solutions based on technology ecosystems. He is an expert in educational vir-tual worlds, data analytics, and technologies for educational purposes, and disciplines that he has developed in many innovation and research projects. He has participated in many European and national projects, such as Tagging, Recognition and Acknowledgment of Informal Learning Experiences, Virtual Alliances for Learning Society, USALSIM Virtual Campus, or the Spanish Observatory for University Employability and Employment, where he participated as a Software Engineer, Researcher, and Developer.

Cristina Maderuelo received the degree in Pharmacy from the University of Salamanca, specializing in Industrial Pharmacy and Pharmaceutical Technology at the University of Barcelona, and PhD from the University of Salamanca. Combines her work as Galenic Technical Developer within the Department of Pharmacy and Pharmaceutical Technology of the University of Salamanca with her teaching activity as Associate Professor at the same university. Cristina is also one of the main promoters of the group Teaching Innovation USALPHARMA, where she directs the teaching activities carried out in the virtual laboratory Usalpharma Lab. She has participated in many projects of pharmaceutical R & D and projects Educational Innovation, publishing the results of both lines of research in many congresses and international magazines.



Francisco José García-Peñalvo received the degrees in computing from the University of Salamanca and the University of Valladolid, and the Ph.D. degree from the University of Salamanca. He is currently the Head of the Research Group Interaction and eLearning. He has led and participated in over 50 research and innovation projects. He was the Vice Chancellor of Innovation with the University of Salamanca from 2007 to 2009. He has authored over 300 articles in international journals and conferences. His main research interests focus on eLearning, computers and education, adaptive systems, Web engineering, semantic Web, and software reuse. He was a Guest Editor of several special issues of international journals, such as *Online Information Review*, *Computers in Human Behavior*, and *Interactive Learning Environments*. He is also a member of the Program Committee of several international conferences and a Reviewer of several international journals. He is currently the Editor-in- Chief of the *International Journal of Information Technology Research* and the *Education in the Knowledge Society Journal*. Besides, he is the Coordinator of the multidisciplinary Ph.D. Programme on Education in the Knowledge Society.

Roberto Therón studied Informatics at the University of Salamanca (Diploma) and the University of La Coruña (BA). After joining the Research Group Robotics at the University of Salamanca, he presented his thesis work, "Parallel calculation configuration space for redundant robots", receiving the Extraordinary Doctoral Award. He subsequently obtained the degrees of Bachelor in Communication Studies (University of Salamanca) and Bachelor in Humanities (University of Salamanca). At the same University of Salamanca he continues to carry out his research work, as manager of VisUsal group (within the Recognized Research Group MIDA) which focuses on the combination of approaches from computer science, statistics, Graphic Design and Information Visualization for obtain an adequate understanding of complex data sets. In recent years, he has been dedicated to developing advanced visualization tools for multidimensional data, such as genetic or paleo-climate data. In the field of Visual Analytics he develops productive collaborations with groups and institutions internationally recognized as the Laboratory of Climate Sciences and the Environment (Paris) or the Center for Advanced Visual Analytics UN (Switzerland). He is the author of over 70 articles in international journals and conferences.

Jonás Samuel Pérez-Blanco holds a degree in Pharmacy and a MSc in Design, Procurement and Drug Evaluation at the University of Salamanca; he currently is an assistant professor of Pharmacy and Pharmaceutical Technology of the University of Salamanca where he is doing a doctoral thesis on monitoring of anticancer drugs. His concentrated trajectory focuses on optimizing pharmacotherapeutic treatments and teaching innovation in the pharmaceutical field. Samuel is the coordinator of new technologies at the Faculty of Pharmacy; He has directed and collaborated on more than 20 research projects and national and international educational innovation such as COPHELA project (Cooperation in Quality Assurance Pharmacy Education and Training between Europe and Latin America). He has also been honored with several awards for his participation in projects of both biomedical character (Academy of Medicine of Valladolid, Spain) and teaching innovation (Academy of Pharmacy of Castilla y Leon, Spain) and has presented the results of his research in journals high impact international as well as national and international conferences.

Hinojal Zazo Gómez is a graduate in Pharmacy and Masters in Design, Preparation and Drugs Evaluation at the University of Salamanca. She is currently a Ph.D. student in pharmacy, within the same university, funded by a grant from the Ministry of Education of the Spanish Government. She simultaneously participates in the teaching innovation group Usalpharma. In 2012 she received the Award of the Academy of Pharmacy of Castile and Leon for her work subsequently published in the *International Jounal of Antimibrobial Agents*. Hinojal has experience in diverse fields such as HPLC analysis, cell cultures or teaching through Virtual Worlds by participating in multiple projects for Research and Innovation, and the subject taught in subjects of the degree of Pharmacy and various courses offered by the University of Salamanca in relation to this area.

Ana Martín Suárez has a PhD in Pharmacy and is a Professor in the Department of Pharmacy and Pharmaceutical Technology at the University of Salamanca. She completed her training as a researcher in the Department of Pharmacy and Pharmaceutical Technology and service of Pharmacy of the University Hospital of Salamanca. She belongs to the GIR of Experimental and Clinical Pharmacokinetics at the University of Salamanca and IBSAL (Institute of Biomedicine at the University of Salamanca). Her main research is in the field of Clinical Pharmacokinetics, addressing drug monitoring in biological fluids and population pharmacokinetic analysis, modeling and simulation PK / PD (Pharmacokinetic / Pharmacodynamic). She has published more than 30 articles, most included in the Journal Citation Reports. With the group of teaching innovation Usalpharma she has developed numerous projects, published articles, conducted training of teachers and received several awards.

6.5 Appendix E. Evolution of the Conversation and Knowledge Acquisition in Social Networks related to a MOOC Course

F. J. García-Peñalvo, J. Cruz-Benito, O. Borrás-Gené and Á. Fidalgo Blanco, "Evolution of the Conversation and Knowledge Acquisition in Social Networks Related to a MOOC Course," in Learning and Collaboration Technologies, vol. 9192, P. Zaphiris and A. Ioannou, Eds. Lecture Notes in Computer Science, pp. 470-481, Switzerland: Springer International Publishing, 2015. doi: 10.1007/978-3-319-20609-7_44.

Evolution of the Conversation and Knowledge Acquisition in Social Networks related to a MOOC Course

Francisco J. García-Peñalvo[1], Juan Cruz-Benito[1], Oriol Borrás-Gené[2], Ángel Fidalgo Blanco[3]

[1]GRIAL Research Group, Department of Computers and Automatics, Research Institute for Educational Sciences. University of Salamanca, Salamanca, Spain
{fgarcia, juancb}@usal.es
[2]GATE, Technical University of Madrid, Madrid, Spain.
oriol.borras@upm.es
[3]Technical University of Madrid, Madrid, Spain
afidalgo@dmami.upm.es

Abstract. This paper presents a real case of tracking conversations and participation in social networks like Twitter and Google+ from students enrolled in a MOOC course. This real case presented is related to a MOOC course developed between January 12 and February 8, 2015, in the iMOOC platform, created as result of the collaboration by Technical University of Madrid, University of Zaragoza and University of Salamanca. The course had more than 400 students and more than 700 interactions (publications, replies, likes, reshares, etc.) retrieved from the social both social networks (about 200 interactions in Twitter and 500 in Google+). This tracking process of students' conversations and students' participation in the social networks allows the MOOC managers and teachers to understand the students' knowledge sharing and knowledge acquisition within the social networks, allowing them to unlock the possibility of use this knowledge in order to enhance the MOOC contents and results, or even close the loop between the students' participation in a MOOC course and the parallel students' usage of social networks to learn, by the combination of both tools using adaptive layers (and other layers like the cooperation or gamification like in the iMOOC platform) in the eLearning platforms, that could lead the students to achieve better results in the Learning process.

Keywords: MOOCs, iMOOC, Conversation, Knowledge Acquisition, Social Networks, Informal Learning, Twitter, Google+

1 Introduction

The informal conversations through social networks are one of the most successful ways to get extra knowledge and enhance the Learning experience in many online courses [1-4]. Many authors have pointed that the conversations and interactions in the social networks could reveal some real characteristics and results of different

Learning activities, offline activities, etc. [5]. For example, it is possible to highlight the theory of Connectivism [6, 7], where the Learning process is enhanced by the connections between students and teachers and online resources [8], refers to the elements of social networks that encourage such relationships or interactions, so these will offer an ideal space for creating Learning communities. According to Siemens [6] from ICTs the importance of individual knows what an shift to what an individual knows how to find out.

According to the literature [3, 9-12], it is possible to identify three types of Learning:

- Formal Learning: "Learning that occurs in an organized and structured environment (in an education or training institution or on-the-job) and is explicitly designated as Learning (in terms of objectives, time or resources). It typically leads to validation and certification".
- Non-formal Learning: "Learning which is embedded in planned activities not explicitly as Learning. Non-formal Learning outcomes may be validated and lead to certification".
- Informal Learning: "Learning resulting from daily activities related to work, family or leisure. It is not organized or structured in terms of objectives, time or Learning support".

In the case of a MOOC, the informal Learning emerges from connections between students in a spontaneous way against non-formal, than even being also a kind of informal Learning is a pedagogical focus directed by the course team, as it would be the case where specific tags in social networks (*hashtags* [13] in Twitter [14] or Google+) are proposed by teachers to start a discussion or conversation out of the MOOC [15, 16]. These kinds of Learning out of the MOOC (informal and non-formal) can influence the results and achievements of the students within the MOOC [17], also it can allow to MOOC teachers, from the study of the conversations of the students, to discover their shortcomings, their major problems faced in the course or even what or how to find solutions to community faced such problems. On the other side it could be a used to reuse in future editions and improve the Learning platform (using the gained knowledge about those subjects that are more interesting for students, those that enhance the informal Learning around the MOOC course, etc.).

The main goal of this research work, based on previous considerations and other that will be discussed below, is to discover the knowledge acquisition and to track the conversations related to the MOOC content courses in environments non-designed for Learning like the social networks [18] to allow MOOC teachers and managers to improve the Learning process in this kind of platforms in future editions of the MOOC courses.

The manuscript is divided into the following sections: Section 1 (Introduction) introduces the problem and main concepts that will be used and discussed in the manuscript. Section 2 (Materials and Methods) presents the resources used to perform the analysis, which are basically the MOOC course, and the social networks where students and course teachers had performed the conversations and where the informal

Learning take place. Section 3 (Results) describes the analysis results, presenting also the data retrieved within the analysis and showing the main trends in conversations, the most used tags for discussions, main users that participated, etc. Section 4 (Discussion) discusses the data presented in the previous section include also considerations regarding also the Learning community features and issues. Finally, the Section 5 (Conclusions) presents several conclusions about the research work and potential work for the future.

2 Materials and Methods

2.1 Materials

iMOOC and the course "Social Networking and Learning"

The intelligent platform MOOC (iMOOC) is the outcome of an agreement of collaboration in 2013 among the Technical University of Madrid, the University of Zaragoza and the University of Salamanca. It is based on the eLearning platform Moodle 2.6.5. Its principal distinguishing features are the adaptability and the promotion of cooperative informal Learning. One of the main features is the use of Cooperative MOOC model proposed by Fidalgo et al. [19] part of a xMOOC (eLearning platforms) combining connectivist characteristics typical of cMOOCs based on Learning communities. In this model, which integrates three layers, we have added a fourth one with gamification elements [20] involving the others. These layers represent the eLearning platform and social networking (technological layer), the instructional design of the course (training layer), the results and generated content from cooperation between students and teachers (cooperative layer). Eventually associated elements are added to the three layers to improve motivation (gamification layer).

Within a single course we find a set of educational itineraries based on three variables: the general user's profile, preferences and choices of the users or students and progress in the Learning process within the course. To configure these features we have used Moodle platform features such as conditional or groups, supported by external plugins to create groupings, obtain statistics or offer certificates automatically. On the other hand the promotion of informal Learning through collaboration has been implemented using tools offered by the platform such as profiles, forums, workshops and external tools like social networks.

The MOOC course "Social Networking and Learning" [21] is an adapted version of the course "Application of social networking to education: virtual communities" version given on the platform Miriada X. The course duration was 1 month, starting on January 12 and ending on February 8, 2015. Regarding the participation, 793 students were enrolled for the course, more than 400 started it, and 183 students finally accomplish the goal. This course aims to teach students to create virtual Learning communities using social networks. Over four modules an overview of the social web, exploring two of the most extended social networks such as Facebook and Twitter is given. In the last module other social networks are studied, without going into too much detail, highlighting those characteristics that define them and can serve for educational purposes. The course takes advantage of features of the platform like is adaptability [22, 23], offering students the possibility of choosing various educational itineraries based on the topics covered in the course. The student can choose from 5 itineraries: Full course for teachers (offers two additional lessons from implementation of Twitter and Facebook to teaching), complete course for non-teachers, Twitter (only one module on the network), Facebook (only one module this network) and special itinerary. The special itinerary was addressed to students who had participated in the course in a previous edition, featuring a new module focused on Learning communities with a more practical approach. This itinerary allowed access to the rest of the course.

***Hahstags* and Social Networks**
Regarding the social part with a non-formal Learning approach, where the connections between the members and the content is generated, it was decided to use a Learning community for the course. To develop the community the course managers chose the Google + and its "Communities" tool, where members could publish using classification categories proposed by the course team. Thereby they could interact, ask questions or discussions and share resources (links, application examples and exercises or activities raised in the course). This community consists of more than 5000 members having students from previous editions of the course and professionals interested in the subject.

To encourage community use throughout the course, different exercises have been proposed to the enrolled students that must be resolved publishing the solutions in this community or in their Twitter accounts, using specific *hashtags* included in the statement of each exercise. Although the discussion of the solution in the social media helped to generate more interaction among students [24]. Both debates as the exercises were associated with a specific category ("discussions" and "activities and exercises"). These relations will also be enhanced off the platform by community discussions or videoconferences a group of students can submit projects related to course topics after voted in the community for the other fellow. These are broadcast live video from YouTube by Google Hangout tool and using *hashtags* students pose their questions to the speakers, both Twitter and from the community in Google +.

As a result of use of *hashtags* within the course, one can distinguish among several types depending on its origin and its use over time:

- Course *Hashtags,* proposed by the teaching staff, would be framed within what is non-formal Learning. There are two types, generic for the entire course or specific for a module, which students could use in their related publications, even in specific activities or exercises for the course.
- *Hashtags* different to those proposed in the course depending on the needs and according to the students' publication. This use would be more associated with informal Learning.

- *Hashtags* used synchronously by participants at specific times of the course, for example *#RSEHangout* for hangouts sessions.
- *Hashtags* proposed in the course and used asynchronously, as the need arises them.

The goal of using these kinds of tags and resources seeks to improve dropout rates due to the heterogeneity of students in the course. Within the MOOC course ecosystem it is also possible to distinguish the three types of Learning explained previously, with the theoretical part in the iMOOC platform corresponding to formal Learning, community Google + created by teachers or *hashtags* proposed by them corresponds to the non-formal Learning and those conversations initiated by students parallel to the course so through social networks or community contributions in periods the course is not taught as informal Learning in the course.

2.2 Methods

To get insight about the usage of social networks and use of conversations and interactions in order to gain knowledge around the MOOC topics, it is necessary to develop a strategy to retrieve, save and use the information that users share on social networks. In this strategy, the authors have developed some crawlers and automatized systems that search in Twitter the usage of some previously-defined *hashtags* related to the MOOC course, determining the amount of users who use it, identifying, if it is possible, the users who are already enrolled in the MOOC course and the users that are only participating in the informal conversation (and without participating in the MOOC), and even those *hashtags* created *ad-hoc* by the students to start new conversations or tag the publications with other extra search. In the case of Google+ the researchers did not develop any crawler due the API restrictions of the social network, but they used third-party tools like AllmyPlus (http://www.allmyplus.com/) that allow them to retrieve social interaction related to each *hashtag*.

The analysis in the first stage was not intended to dig inside the real conversations (was not intended to make text mining or other text analysis techniques), but it intend to reveal the use of some *hashtags*, the interaction among users, etc. that will serve as basis for later analysis that could reveal the URLs that users share, the identification of leaders and influencers in the conversations, determining coincidences between the students or users of the social network, their usage rates of some resources, etc. in order to get a full insight about the MOOC community that could be used to improve it for later editions of courses. This full insight will reveal some aspects like the students' interaction and conversations, and what kind of informal or non-formal Learning happens in detail in these social networks (which are not specialized in academic content but intended for general purpose).

Once the analysts have retrieved and saved the information, it is needed to define how can be possible to extract true knowledge from the raw data, and how the system could reveal and show this knowledge to the analyst, MOOC managers, etc. In this first approach of the analysis, the representation is performed mainly by through tables and structured data, also some basic graphs were implemented to help the analysts to understand the information presented.

3 Results

The application of the previously described information tracking strategy in the social networks led the authors to discover some relevant information. This information help to understand the scope of the social interactions between MOOC course users, and MOOC course contents. As previously stated, the course teachers proposed several *hashtags* to help the conversation tagging and tracing its evolution, or simply to tag the common conversations about the MOOC contents (*hashtags* related to non-formal Learning). These *hashtags* were the following: *#DebatesRSE, #ActividadesRSE, #DudasRSE, #AvisosRSE, #EjerciciosRSE, #modulo1RSE, #modulo2RSE, #modulo3RSE, #modulo4RSE, #RSEMOOC, #RSEHangout, #RSEEjemplosRRSS, #RSEMalasPracticas, #RSEmiKlout, #UsosTwitterEnseñanza, #RSEMoodleTwitter*.

The analysis of these *hashtags* revealed the following amount of interactions (publications, replies and comments to the publications, retweets and reshares of publications, and favorites or +1 in Google+) and its distribution over both social networks (Table 1, Figure 1)

Table 1. Total interactions in Twitter and Google+ with teachers' proposed *hashtags*

Total interactions	*Twitter*	*Google+*	*Total*
Publications	108	119	227
Replies/Comments	17	76	93
Retweets/Reshares	42	17	59
Favorite / +1	45	315	360
Total	212	527	739

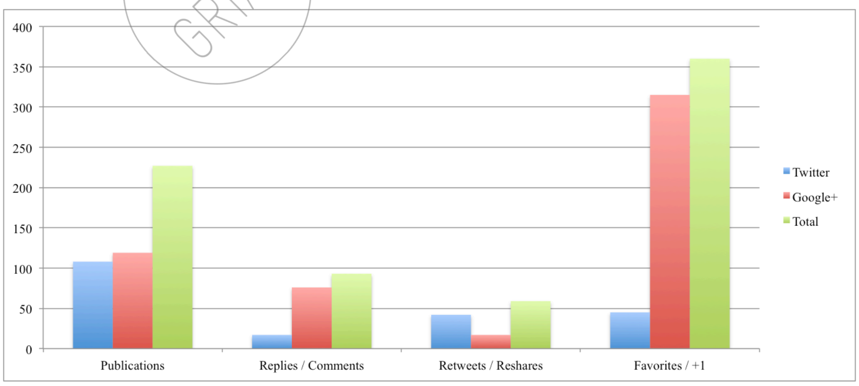


Fig. 1. Total interactions versus interactions in each social network related to the MOOC course

Among these global data per social network, it is possible to filter the information depending the type of interaction and each *hashtag* so it is possible to know the real interactions with the contents related to the MOOC course (through its tagging by *hashtags*). Following are presented the data of the most used *hashtags* proposed by teachers (Table 2, Figure 2)

Table 2. Official *hashtags* interactions in each social network (*hashtags* most used)

Interactions/ Hashtag	*#RSEMOOC*	*#RSEHangout*	*#RSEEjemplosRRSS*	*#RSE-MalasPracticas*	*#RSEmiKlout*	*#RSE-MoodleTwitter*	*Total interactions per type*
Twitter Tweets	9	19	4	5	8	59	104
Google+ Publications	16	4	35	27	20	0	102
Twitter Replies	2	4	1	0	1	9	17
Google+ Comments	33	15	9	2	8	0	67
Twitter Retweets	5	16	0	1	5	9	36
Google+ Reshares	3	2	6	5	1	0	17
Twitter Favorites	5	15	0	2	6	11	39
Google+ +1's	57	25	84	47	51	0	264
Total Hashtag Interactions	130	100	139	89	100	88	

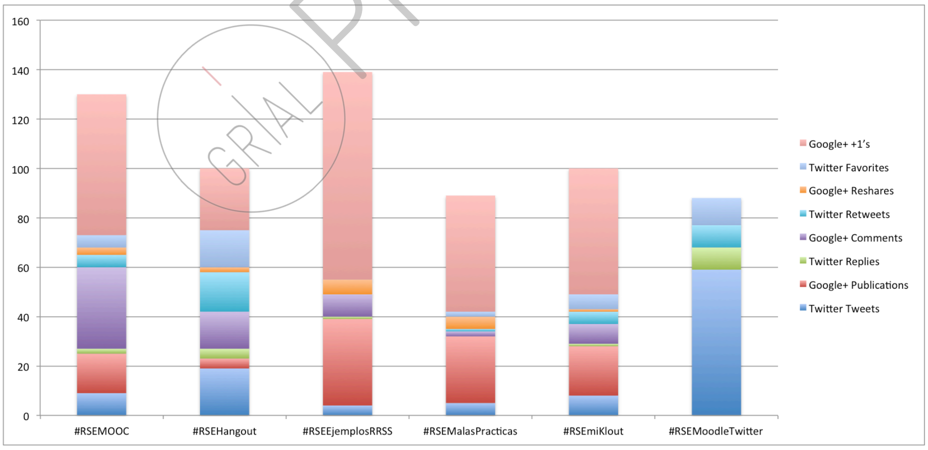


Fig. 2. Distribution of interactions in each proposed *official hashtag* in the social networks

Also, the teachers proposed some debates (non-formal Learning) in the social networks (mainly in Google+) to discuss some concepts near to the MOOC contents. In the case of this course, the teachers used one debate started in the previous course, obtaining in this edition 28 comments and 5 +1's on Google+.

Table 3. Unofficial *hashtags* most used by the students within the MOOC course (related to informal Learning) in Google+

	#facebookeducacionarse	*#twitter*	*#educación*	*#facebook*	*#aprendizaje*	*#infografía*	*#aula*
Number of publications	26	12	10	9	7	6	4
Comments	7	18	14	7	4	13	0
Reshares	5	21	20	6	8	7	0
+1's	61	94	65	54	55	35	4
Total Interactions / Hastags	99	145	109	76	74	61	8

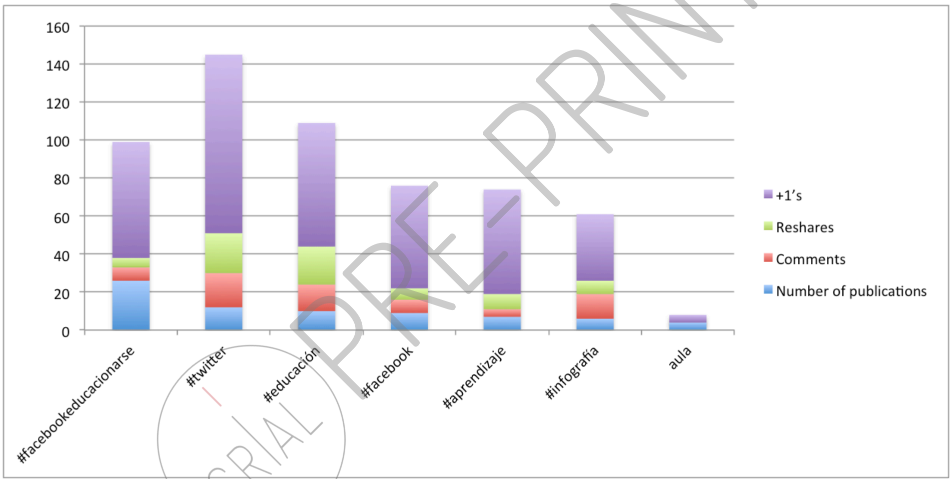


Fig. 3. Distribution of interaction with unofficial most used *hashtags* proposed by students in Google+

Regarding the informal Learning component of the social networks usage, the analysis revealed some interesting data about the trends in informal conversations and learners' knowledge sharing preferences. These data, as the previously explanation about the teachers' proposed *hashtags*, can be showed filtered by social networks interactions with *hashtags* and debates started (by the students in this case). As example, in the case of *hashtags* used by students, not proposed by the teaching staff, it is possible to find, in Google+ 169 publications started by students. In these publications, the *tags* most used (and most interacted) by them are the following (Table 3, Figure 3): *#facebookeducacionrse*, *#twitter*, *#educación* (education in Spanish language), *#facebook*, *#aprendizaje* (Learning in Spanish), *#infografía* (infographics in Spanish), *#aula* (classroom in Spanish).

Also, about *unofficial* and *informal* debates started by the students, the analysis raised 10 different discussions with 14 publications and 45 +1's also on Google+ community.

4 Discussion

Regarding the results, the authors want to discuss two main questions about the utility of the study and its application. These questions are: Is it possible to identify the *trending topics* (subjects that interested more to the MOOC and social networks audience) in this Learning community through this analysis approach? Is it useful to know these data to improve the MOOC or the Learning community?

From the authors' point of view, the answer to both questions is yes:

- Yes, it is possible to identify in the social networks the trending concepts and contents most discussed and probably more interesting for the students in both ways, through *official hashtags* and through *unofficial hashtags*. Regarding the data, it is possible to identify what are the most used tags, what contents reflect more users' interactions, etc., so it is possible to assert that retrieving information about the interaction with the tags, the teachers and MOOC managers could identify what subjects and concepts are more interesting for the users and apply this knowledge to improve the experience within the MOOC and its results. Also, it should be remarked that if this analysis approach is enriched, for example with basic text mining techniques, the teachers and managers could filter the interactions and comments using extra information that learners use, like URL, news posted, content feeds, etc., and this can open new advanced possibilities even helping the personalization and adaptation to users that iMOOC approach performs.
- Yes, the data is useful to improve the MOOC and enhance the Learning community. The retrieved knowledge can be used, as stated just before, to "close the loop" with the adaptativity, cooperation or gamification features present in the iMOOC platform, or even closing the loop in other way, using both kind of tools (MOOC and social networks) to establish tools collaboration approaches to use them within the same Learning processes. For example, knowing the main interests for a user (based on the comments and interactions inside and outside the MOOC platform), the iMOOC could present contents or resources depending on the user's interests detected within the social networks, or even MOOC could recommend some debates and conversations in the social networks, etc. based on the users' interaction inside the MOOC course.

There are another two issues that authors would remark regarding the conversations tracking; one is the limitation retrieving data from the social networks, and the other main issue is related to the previous students' skills using social networks and the errors and mistakes they can make using the systems, the *hashtags*, etc.

In the case of the limitations retrieving information, there are many problems with Twitter and Google+.

- In the case of Twitter, the API only allows to search tweets in a 7-day window before the moment of the search, so it difficult too much to perform the analysis post-course. Instead of this search methods, the Twitter's API allows to *live stream* the tweets under some *hashtags*, but it requires that the analysts know previously all the possible *hashtags* that would be used, or another techniques that make possible to include new *hashtags* within the live tweet stream.
- In the case of Google+, Google APIs limits the access to retrieve data, so if analysts want to retrieve the interactions and data about the activity on Google+ without restrictions, they should perform manual analysis tasks, use web scrapping techniques, or utilize third-party tools like those used in this research work (http://www.allmyplus.com/ for example).

Other relevant issue regarded in the analysis, is the importance of the students' previous skills using these kinds of systems like the social networks. During the analysis, the researchers have observed many errors using tags in publications, error commenting other activities performed purely in the social networks, etc. These skills and the performed mistakes are relevant in the analysis because they could introduce noise and errors in the results. Thus, the researchers would develop strategies in the future to avoid this kind of noise and possible errors during the analysis phase.

5 Conclusions

This paper explains the authors' way to track the conversations in social networks related to a MOOC course and how they make basic analysis in order to review the knowledge sharing and knowledge acquisition through these social networks. The paper also reveals how this kind of data retrieval and basic analysis empowers the MOOC managers and teachers to understand and track the conversations in social networks like Twitter and Google+ around MOOC concepts and subjects, so they can measure the Learning process in these social networks regarding three kinds of Learning, formal, informal and non-formal Learning. The data presented in this paper can serve as a basic example of this usage, and the authors present some other applications and future work to enhance and improve this analysis to make it more powerful.
Regarding the results of this preliminary study about the students' conversations in social networks and its reflection on their knowledge acquisition and Learning, the authors agree that it is possible to enhance the retrieval and analytics process, achieving a clear process with great outcomes regarding the detection of interests, desires and concepts and subjects that the students want to discuss with others.

The improvement of this process, would allow achieve the combination of both Learning tools, the MOOC platform or eLearning platform that the students use to learn in a formal or non-formal ways, and the social networks that the students use to learn in non-formal or informal ways. This combination could produce a loop process where the MOOC and social networks can feedback themselves, detecting students' behaviors, desires and interest, to use them later by the integration with adaptive Learning platforms, like the iMOOC platform or many others, improving by this way the Learning processes through the use of personalization layers that use interests detected and the insights retrieved from users' interaction in social networks, to present personalized contents to the students that could encourage them to improve their Learning, to obtain better outcomes from this process, and better performance in the Learning experience.

6 Acknowledgments

We'd like to thank to the GRIAL and LITI research groups by their support in this research and also we'd like to thank to the UPM by its financial support throughout the "Diseño y desarrollo de MOOC universitarios" (Ref. PT1415-05000) Project.

7 References

1. Mackness, J., Mak, S., Williams, R.: The ideals and reality of participating in a MOOC. 7th International Conference on Networked Learning, pp. 266-275 (2010)
2. McAuley, A., Stewart, B., Siemens, G., Cormier, D.: The MOOC model for digital practice. SSHRC Knowledge Synthesis Grant on the Digital Economy (2010)
3. Vanbaelen, R., Harrison, J., van Dongen, G.: Lifelong learning in a Fourth World setting. Professional Communication Conference (IPCC), 2014 IEEE International, pp. 1-9 (2014)
4. García-Peñalvo, F.J., Johnson, M., Ribeiro Alves, G., Minovic, M., Conde-González, M.Á.: Informal learning recognition through a cloud ecosystem. Future Generation Computer Systems 32, 282-294 (2014)
5. Aramo-Immonen, H., Jussila, J., Huhtamäki, J.: Visualizing informal learning behavior from conference participants Twitter data. Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality, pp. 603-610. ACM, Salamanca, Spain (2014)
6. Siemens, G.: Connectivism: A learning theory for the digital age. International journal of instructional technology and distance learning 2, 3-10 (2005)
7. Zapata-Ros, M.: Teorías y modelos sobre el aprendizaje en entornos conectados y ubicuos. Education in the Knowledge Society 16, (2015)
8. Evans, C.: Twitter for teaching: Can social media be used to enhance the process of learning? British Journal of Educational Technology 45, 902-915 (2014)
9. Marsick, V.J., Watkins, K.E.: Informal and Incidental Learning. New Directions for Adult and Continuing Education 2001, 25-34 (2001)
10. García-Peñalvo, F.J., García-Holgado, A., Cruz-Benito, J.: Formal and informal learning experiences in multicultural scopes. Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality, pp. 523-527. ACM, Salamanca, Spain (2013)

11. Bjornavold, J., Training, E.C.f.t.D.o.V.: Validation of non-formal and informal learning in Europe: a snapshot 2007. Office for Official Publications of the European Communities, Luxembourg (2008)
12. García-Peñalvo, F.J., Colomo-Palacios, R., Lytras, M.D.: Informal learning in work environments: training with the Social Web in the workplace. Behaviour & Information Technology 31, 753-755 (2012)
13. Ghenname, M., Abik, M., Ajhoun, R., Subercaze, J., Gravier, C., Laforest, F.: Personalized Recommendation Based Hashtags on E-learning Systems. ISKO-Maghreb'2013 Concepts and Tools for Knowledge Management (KM), pp. NA, Tunisia (2013)
14. Twitter Inc,, https://support.twitter.com/articles/49309
15. West, J.: Recognition of non-formal and informal learning: the Case Against. Study prepared for the meeting of the OECD Group of Experts. (2007)
16. Vosecky, J., Jiang, D., Leung, K.W.-T., Xing, K., Ng, W.: Integrating Social and Auxiliary Semantics for Multifaceted Topic Modeling in Twitter. ACM Trans. Internet Technol. 14, 1-24 (2014)
17. Fidalgo-Blanco, Á., Sein-Echaluce, M.L., García-Peñalvo, F.J., Esteban Escaño, J.: Improving the MOOC learning outcomes throughout informal learning activities. Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality, pp. 611-617. ACM, Salamanca, Spain (2014)
18. Alario-Hoyos, C., Pérez-Sanagustín, M., Delgado-Kloos, C., Parada G, H., Muñoz-Organero, M., Rodríguez-de-las-Heras, A.: Analysing the Impact of Built-In and External Social Tools in a MOOC on Educational Technologies. In: Hernández-Leo, D., Ley, T., Klamma, R., Harrer, A. (eds.) Scaling up Learning for Sustained Impact, vol. 8095, pp. 5-18. Springer Berlin Heidelberg (2013)
19. Fidalgo, Á., Sein-Echaluce Lacleta, M.L., García-Peñalvo, F.J.: MOOC cooperativo. Una integración entre cMOOC y xMOOC. In: Fidalgo Blanco, Á., Sein-Echaluce Lacleta, M.L. (eds.) Actas del II Congreso Internacional sobre Aprendizaje, Innovación y Competitividad, CINAIC 2013 (Madrid, 6-8 de noviembre de 2013), pp. 481-486. Fundación General de la Universidad Politécnica de Madrid, Madrid, España (2013)
20. Borrás Gené, O., Martínez Núñez, M., Fidalgo Blanco, Á.: Gamification in MOOC: challenges, opportunities and proposals for advancing MOOC model. Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality, pp. 215-220. ACM, Salamanca, Spain (2014)
21. Technical University of Madrid (Spain), University of Zaragoza (Spain), http://gridlab.upm.es/imooc/course/view.php?id=2
22. Fidalgo Blanco, Á., Sein-Echaluce Lacleta, M.L., García-Peñalvo, F.J.: Methodological Approach and Technological Framework to break the current limitations of MOOC model. Journal of Universal Computer Science In Press, (2015)
23. Fidalgo Blanco, Á., García-Peñalvo, F.J., Sein-Echaluce Lacleta, M.L.: A methodology proposal for developing adaptive cMOOC. In: García-Peñalvo, F.J. (ed.) Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality (TEEM'13), pp. 553-558. ACM, New York, NY, USA (2013)
24. Karaoglan, B., Candemir, C., Haytaoglu, E., Algin, G.B., Demirci, S.: Using Twitter as a diagnostic teaching and learning assessment tool. EAEEIE (EAEEIE), 2014 25th Annual Conference, pp. 73-76 (2014)

6.6 Appendix F. Extending MOOC ecosystems using web services and software architectures

Extending MOOC ecosystems using web services and software architectures

Juan Cruz-Benito
GRIAL Research Group
Department of Computer and Automatics, University of Salamanca
IUCE, Paseo de Canalejas 169
37008, Salamanca, Spain
juancb@usal.es

Oriol Borrás-Gené
GATE, Technical University of Madrid
Calle Ramiro de Maeztu. 7
1st floor (Rectorado A)
28040 Madrid, Spain
oriol.borras@upm.es

Francisco J. García-Peñalvo
GRIAL Research Group, Department of Computer and Automatics
Research Institute for Educational Sciences (IUCE), University of Salamanca
Paseo de Canalejas 169
37008, Salamanca, Spain
fgarcia@usal.es

Ángel Fidalgo Blanco
Technical University of Madrid
Calle Alenza nº4, office 718
(M Building) 7th floor
28003, Madrid, Spain
afidalgo@dmami.upm.es

Roberto Therón
Department of Computer and Automatics
Faculty of Sciences, University of Salamanca
Plaza de los Caídos s/n
37008, Salamanca, Spain
theron@usal.es

ABSTRACT
This paper present a research project that tries to extend the MOOC ecosystems by integrating external tools like social networks. This integration is developed by using a software architecture that mediate between the different systems and platforms establishing communication workflows and analyzing the information retrieved. This kind of system is applied in a real case, and it allows teachers and managers of the MOOC platform to get enhanced information and insights about users interaction with contents and MOOC tools, as well as some metrics impossible to retrieve or calculate manually in this kind of eLearning platforms with high amounts of users.

Categories and Subject Descriptors
H.3.5 **[Online Information Systems]**: Web-based systems. H.5 **[Information Interfaces and Presentation]**: General K.3.1 **[Computers Uses in Education]**: Collaborative learning.

General Terms
Measurement, Human Factors.

Keywords
iMOOC, MOOC, Moodle, Web Services, Software Architectures, HCI, eLearning.

1. INTRODUCTION
The analysis of interaction among users and systems provide great insights about how users use, understand and take advantage of tools and platforms they utilize to perform any kind of task.

The fact of analyze the interaction and try to extract valuable knowledge from it, have real application in many areas of knowledge and business, as in digital marketing, in education (Learning Analytics), etc. Although in some fields as education, this type of behavior analysis, and interaction analysis is increasingly common, the approaches and tools developed should be updated and adapted to new systems, platforms and paradigms in eLearning. In these new types of learning platforms and paradigms can be highlighted the MOOCs, because they expand the traditional limits in students' interaction with teachers, contents and online learning platforms. Furthermore, MOOCs leverage other platforms (even those that are not purely intended to be applied in education) like the social networks and other online tools, applying by this way multi-platform and multi-context approaches that can improve and upgrade the learning experience [20].

It is because of this use of multiple tools and multiple context that is necessary to design and implement new ways of interaction analysis and platforms that allow to perform it. These new ways and platforms will manage the acquisition of knowledge regarding the learning and interaction with platforms, establishing convergence of knowledge between different learning vectors and context, to finally allow teachers and managers to learn, explore and implement possible improvements that help in the learning process, the design of content and the motivation of students.

The goal of this paper is to explain a modular software architecture implemented to allow teachers and managers of a MOOC retrieve knowledge about how users enrolled in a MOOC course utilize some tools external to the MOOC platform, getting by this way insights about what did users on these external tools, what kind of interaction they perform inside them, and thus, discover possible improvements and solutions for eLearning processes to be applied later in the MOOC platform and its courses.

To explain these contents, the paper is divided into the following sections: section 2 (Aims and goals) presents the main aims and goals of the research presented. Section 3 (iMooc Platform) describes real MOOC platform where is being tested the software architecture proposed and the integration with other online tools. Section 4 (Software Architecture Proposal) explains the software architecture designed to tackle the integration with other tools and gain knowledge about the MOOC users and their interaction with contents inside or outside the MOOC platform. Section 5 (Services and Crawlers) explains how is being developed the software architecture integration with other tools and platforms (eLearning platform and social networks). Section 6 (Results) shows partial results of the application of this software architecture retrieving data from the integrated environments. Finally, the Section 7 (Conclusions) presents several conclusions about the research work and potential work for the future

2. Aims and goals

The MOOC courses are characterized by a high drop out rates primarily due to the heterogeneity of participants is why further information on their activity on the platform will assist meet the shortcomings of course adapted to the student needs.

It aims to create a model to obtain the maximum information from the MOOC platform combining it with the results obtained in the external elements that support the learning of the course, such as social networks, to give a feedback to the educational process, analyze performing users to adapt the content to their interaction, study the behavior to provide the extra functionality MOOC integrating user feedback analysis.

Thus the analysis of interaction in media is emphasized to obtain a deeper comprehension of users or the use of content analysis possible interest. Detecting those topics that interest students while those resources contributed by students for future editions are reused.

It is important to locate the types of learning according to the proposed activities. It is in the non-formal and informal learning where these foster interactions between users

Because of the massiveness of these many interactions resources and resources making it necessary to use tools to filter and collect all the information generated to recover this. This paper is based on the use of hashtags associated with publications to quickly retrieve the generated content. Hashtag is a character string with a topic associated preceded by a # sign and is used to mark tweets [14], part of the text of these. Although their use is widespread in twitter have extended to other social networks such as Google+. A study of hashtags is posed for the types of learning identified in the elements and course activities.

3. iMOOC Platform

3.1 The iMOOC approach

MOOC courses offer new opportunities for learning, features like massiveness of participants, peer-to-peer interactions, free-of-charge, openness or scalability [28]. There are two main types of courses MOOC, the xMOOC with a behavioral approach (occurs in traditional online courses) and cMOOC with a connectivism and networking based approach [17]. Taking advantage of features of both types of MOOC can apply the cooperative model of Fidalgo et. al. [18] which can be defined by three layers: The first one is the “technological layer” linked to the platform where will find the course content and the social network that will support the learning community. The second layer named "training strategy" refers to the instructional design of the course. This layer is divided, according to the two types of MOOC exposed, into a “behavioral strategy” focused on the acquisition of basic common knowledge and into a “connectivist strategy” dedicated to the generation of educational resources by participants. At this point, the resources generated from both strategies need organized. Finally, the “cooperative layer” shows the outcomes and the content generated with the cooperation of instructors and participants of the course [18]. For more connectivist orientation course based on the interaction between participants will require the use of social tools, a good selection of these is a good way to engage students and promote their participation in the course [1] using these tools we can create virtual learning communities (VLC), it is defined as communities of people who share common values and interests, and that are connect via different communication tools that such networks offer, whether synchronous or asynchronous [5]. The incorporation of VLC may provide greater interaction between participants, support and guidance to people with difficulties and may increase collaborative processes between participants.

Over the life of people not only learn in schools or Universities, so do in other situations in life, this idea is called lifelong learning MOOC one of the developments associated with this item. Based on the concept of lifelong learning are identified three types of learning associated with this concept [16]. Their differences [15] are based on the following items: where does the learning take place, in the case of online learning, the platform becomes as the institutional place for learning; if the learning is structured or unstructured; whether the learning is guided or not by faculty; and if there is any certification [2].

Taking these differences the Formal learning is that occurs in an organized and structured environment and leads to validation and certification, it is the most institutionalized [2]. The following types of learning take place in less organized spaces. Then, with opposite sense to the formal is informal learning, which is voluntary and unstructured. It is based on the intrinsic motivation [11] of the student. In this type of learning, the student chooses the way to acquire the knowledge, the learning is given everywhere and applies to any situations in common life [15] and is not evaluated. Finally there is the non-formal learning, which is structured and guided by the faculty, but is generally voluntary and is usually not evaluated. MOOC elements can be associated with the three types of learning, finding the more formal part in the course platform (xMOOC) and the informal part and non-formal community-associated learning (cMOOC).

3.2 Description and main features of the platform

In 2013, because of the agreement between the Technical University of Madrid, the University of Zaragoza and the University of Salamanca, the platform iMOOC or intelligent MOOC [32] emerges. Based on Moodle 2.6.5 platform [27] was chosen for its versatility. The main MOOC platforms such as edX [13], Coursera [6] or MiriadaX [24] focused on Spanish speaking require a single path for the course. Given the heterogeneity of the participants this can be in many cases the abandonment of them to take the part that interests them, one of the main distinguishing features of the iMOOC platform in front those with a more traditional approach is adaptivity of courses for students. This adaptivity is based on three variables: Depending on the user profile, according to an itinerary chosen by the student or for student progress within the course and the knowledge that he is gained. To achieve this functionality is necessary to use Moodle features such as conditionals and groups, as well as external plugins to create groupings. On the other hand iMOOC provides an appropriate environment to an Informal and cooperative learning.

Will make use of the tools offered by the platform such as profiles, forums, workshops and secondly external social networks for informal.

3.3 MOOC course "Social Networking and Learning"

To take advantages of iMOOC platform in early 2015 a course based on a test offered by the faculty version along two editions in the Iberoamerican MOOCs platform Miriada X [4] was launched. The course of one-month duration began January 12 and ended on February 8, leaving an extra week to allow students to finish it. A Total of 793 students were enrolled for the course, more than 400 started it, and 183 students finally accomplish the goal.

The VLC was created using the tool communities of Google +, space where students can interact with each other and with content in the network sharing it. Here they can also discuss, submit questions and publish the result of voluntary exercises throughout the course. To do this has been created 9 default categories (presentations, announcements, discussions, questions, resources, Activities and exercises, application examples, contests and more) with which to classify each publication, besides the possibility to use hashtags.

This course is based on the cooperative model [18] at which has been added a fourth layer called "Gamification layer" that interacts with the other three layers, which aims to improve motivation [3]. Depending on types of learning can be divided the course study into three parts, the first section concerns the iMOOC platform itself, which takes place the more formal or formal part of the course and the other two sections relate to the community of external learning the course where interactions among participants are established and where they are generated and share new resources.

The objective of this course is to provide participants, mostly teachers of basic digital skills in social networks to implement in their classrooms with students creating virtual learning communities. Throughout 4 modules consist of lessons, it is theoretically analyzed in depth the social web and learning communities are studied thoroughly Facebook and Twitter networks and finally reviewed with less detail other valid networks for educational use with and support tools to manage social networks. The modules are divided into small lessons that address a specific topic and the content is based on a short video as well as additional information that complements the video.

As adaptive part participants can choose between five educational itineraries: Full course for teachers (additional lessons are given the keys to implement networks studied), complete course for non-teachers (only the different networks are studied), Twitter (only one module on the network), Facebook (only one module this network) and one special itinerary. The special itinerary was addressed to students who had participated in any of the two previous editions, with an extra module focused on learning communities.

The course consists of a series of directed activities and proposals by the faculty to enhance knowledge of the lessons and interaction between course participants:

- Voluntary exercises along the videos. In both cases this activity takes place in the learning community.
- Discussion proposed and discussed in the community.
- Two videoconferences using Google Hangouts, streamed live, where some course participants could make a brief presentation to other peers about an educational project related to his social networks

Table 1. Relationship of course activities on Google +

Activity	Category	Hashtag
Search examples of social networks	Activities and exercises	#RSEEjemplosRRSS
Bad practices in the use of social networks	Activities and exercises	#RSEMalasPracticas
Measuring influence in social networks using Klout	Activities and exercises	#RSEMiKlout
Using Twitter in education	Activities and exercises	#UsosTwitterEnseñanza
Discussion about the possibility of replacing a learning management system (LMS) through a social network	Discussions	
Discussion on digital identity	Discussions	
Hangout		#RSEHangout

These activities can also be performed on Twitter although the official platform is Google +. Course evaluation is based on four questionnaires one per module that participants must overcome to obtain a certificate of participation offered for free.

So this part can be associated to a non-formal learning, since such courses by not offering an official certifications, for example to allow for recognition of credits, can not be considered as formal learning.

About informal learning this occurs in the community of Google + and is associated with the interactions among participants not addressed by the teaching staff, similar to learning which could result in informal settings outside the classroom such as between colleagues in an office. Three types of interactions:

- Creating new hashtags and subsequent interaction
- Discussion proposed by participants
- Resources contributed

The last two actions are included in the categories created for this publication ("debates" and "resources").

4. Architectural proposal

Following the previous experience of the authors in similar cases, where they apply software architectures to extend the functionality of eLearning ecosystems [7-10; 21], authors decided to use the core of a software architecture they built in 2014.

Several layers compose this core, one to retrieve data from each external platform or tool, other that wipes and stores the information retrieved, another to push analyzed information to other platforms, and others that enable searches and interaction between information and users.

The core of the architecture is a system that acts as a mediator between the different social networks and learning platforms that will be interconnected (Figure 1). This mediator system communicates with each external tool through using web services (REST APIs commonly) and crawlers; retrieving data and information from them and analyzing the information in order to convert the raw data in valuable information for teachers and managers of the iMOOC platform (based on Moodle).

In order to implement this software architecture and its layers, has been used for development several technological components and technologies that are listed below:

- Django Framework [12]: This web framework is used to build the software layers and to coordinate the information workflows between the components and systems of the architecture.
- MongoDB [25]: This NoSQL database is used to store the data without the traditional restrictions of the SQL databases, and allows to adapt the database storage schemas to each kind of content retrieved from external tools and platforms [7].
- REST APIs: these web services are used to serve as communication channels between components and systems involved in the MOOC ecosystem. In case of those tools and systems that do not provide this kind of facilities, will be used crawlers to retrieve information (this will be explained in the following section).

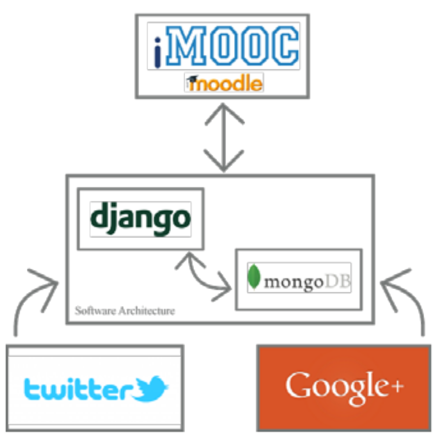


Figure 1. Software architecture proposal

The main idea behind the framework, is that teachers and managers could use the website provided by the architecture to interact with the information and data retrieved from the external tools, so all the assessment and evaluation of the users learning in the MOOC could be centralized in the architecture. The other possible approach, is that the architecture push the analyzed information again to the Moodle platform, allowing teachers and managers to allow them assess and evaluate the learning without leave the MOOC platform, in that case, the data visualization and interaction with the information retrieved depends strictly on the resources Moodle provides.

5. Services and crawlers

In order to implement the information workflows shown in the Figure 1 between the software architecture and the different systems and social networks, researchers need to establish the proper communication channels for the information. These communication channels are based, in this case, in services and crawlers:

- The services are facilities provided by third-party software to facilitate the communication and interconnection with other systems, applications or clients. In this case, researchers have used services for communicating with Moodle and Twitter.
- The crawlers are software applications that find automatically information in third-party systems when they do not provide services for *pull* and *push* information between systems. In this case we are working on crawlers for getting information from Google+ Communities (Google+ does not provide API or other services to get and post activities and other information within the communities).

In the subsections below are explained how have been used these services and crawlers, and how they are implemented within the software architecture.

5.1 Moodle

Moodle provides several API services and API architectures; allowing users and third party applications and systems interact with courses, administration settings, users and configuration information. The API used in this work is based on Representational State Transfer (REST) architecture [19], and it allows several actions in both directions of communication (GET and POST actions, as well as DELETEs, etc.); the full documentation and functionalities of this API can be found at [26].

For example, these API endpoints and functionality allow managers and teachers of the (i)MOOC course to make automatic checks about the tasks completed by users, automatized (and simply) assessment about their participation in the MOOC, etc. In a regular course on Moodle, this usage of the API is not a key aspect, most of these checks and assessment is performed manually by the teachers, but in a MOOC course with more than 700 hundreds (in this case, several thousands in bigger MOOC courses) turns out this resources as a key factor to evaluate the users' interaction with the MOOC and for assessing their learning within the course.

Bellow is presented an example code that allow teachers to retrieve the full list of users enrolled in the course; this result list of users enrolled, for example in the case of the iMOOC course was used to check what users filled their profile with the links to their personal social networks profiles, which was proposed as an activity of the *Twitter in education* lesson. As previously explained, authors implemented the software architecture using Python language, so the code is formatted in the *phytonic* way and includes the main software library used, Requests [30] that allows to implement easily the API consumption.

```
import requests, json
```

```
parameters = {'wsfunction': core_enrol_get_enrolled_users', 'courseid':'id',
              'moodlewsrestformat':'json', 'wstoken':'xxxxxx'}
url = "http://gridlab.upm.es/imooc/"

response = requests.get(url, params=parameters)

if response.status_code == 200:
   results = response.json()
   for result in results:
      print result
else:
   print "Error code %s" % response.status_code
```

5.2 Twitter

Regarding the Twitter data retrieval implementation, the authors have implemented collector that gets tweets on live based on their hashtags. This implementation is possible thanks to the Twitter REST APIs [31] and Tweepy library for Python [29]. Using both facilities (specially the Twitter Streaming API) authors built a software that is able to retrieve in real time tweets tagged [23] with the any of hastaghs proposed in the MOOC course and storing the tweets in the software architecture database (enabling by this way MOOC user matching, etc). As example of how is done this data retrieval, below is attached a simplified version of the code:

```
from __future__ import absolute_import, print_function
from tweepy.streaming import StreamListener
from tweepy import OAuthHandler
from tweepy import Stream

consumer_key="xxxx"
consumer_secret="xxxx"
access_token="xxxxx"
access_token_secret="xxxxx"

class StdOutListener(StreamListener):
   def on_data(self, data):
      try:
         print(data)
         return True
      except:
         pass
   def on_timeout(self):
      sys.stderr.write("Timeout, sleeping for 60 seconds...\n")
      time.sleep(60)
      return

   if __name__ == '__main__':
      l = StdOutListener()
      auth = OAuthHandler(consumer_key, consumer_secret)
      auth.set_access_token(access_token, access_token_secret)
      stream = Stream(auth, l)
      stream.filter(track=['#RSEEjemplosRRSS',#UsosTwitterEnseñanza",
                           "#RSEMiKlout "])
```

About Twitter integration in the system, should be highlighted that the MOOC managers and teachers must get permission of the users about storage their tweets, or simply anonymize the personal data present in each tweet (name and username, location, etc.), because the social network specify in their policy rules this restriction.

5.3 Google+

About Google+, the situation is totally different. This social network provides APIs and methods to retrieve information about users, posts, comments, etc. [22], but it does not allow to retrieve information from the users communities within the social network. This disables the possibility of use the same way to get information about conversation and interactions in the communities, regarding this, teachers and managers from the MOOC course were searching other tools that let them to retrieve the desired information; for example, they use currently the tool AllMyPlus (http://www.allmyplus.com/) that allows them to retrieve information of the learners community related to MOOC. This is not the best solution, because it convert the ideal automatic process indeed in a manual process, so the authors are trying to develop a crawler that enables them to retrieve information directly from Google+ website or AllMyPlus website.

6. RESULTS

By using the software architecture and the other tools (AllMyPlus), was possible to retrieve information about users' posts on social networks, information about their profiles on the iMOOC platform, etc. As example below of the utility of this kind of software architecture supporting and expanding MOOC functionalities, are shown several metrics retrieved from the complete learning ecosystem:

- During the course were recorded in the Google+ community 302 publications belonging to 140 users, 57 of whom have used a hashtag. Table 2 shows the times that have been used hashtags and how many of them are depending on the type of learning.
- During the course were retrieved more than 200 interactions of MOOC users with the content and hashtags in the social network Twitter.
- Also, to evaluate the completion of the MOOC activities, for example as mentioned in previous section, related to the number of users that fill their MOOC profiles with the links to their social networks profiles. In this case, the number of users that accomplish this activity was 275, a 33'86% of the total users.
- Hashtags usage and temporal evolution (Figure 2) of proposed hashtags by the teachers (non-formal) and those used by students initiative (informal).

Table 2. Distribution of posts and contents in Google+ community

	Total #	Different #	# misspelled	Users using #
Non-formal	128	8	8	37
Informal	144	82	-	43
Total	272	90	-	23 (both types)

Table 3. Interactions with MOOC contents and proposed hashtags in Twitter

Total interactions	**Twitter**
Publications	108
Replies	17
Retweets	42
Favorites	45

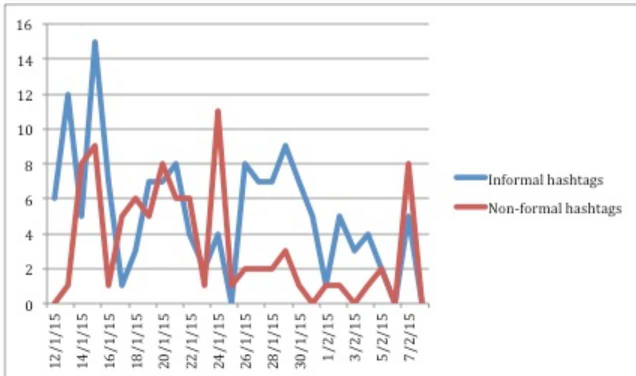


Figure 2. Evolution of informal and non-formal hashtags related to MOOC contents usage

7. CONCLUSIONS

This paper explains a software architecture designed and developed to extend the MOOC ecosystems functionalities and utilities by integrating external tools like social networks. This integration is built by using a software architecture that mediate between the different systems and platforms establishing communication workflows and analyzing the information retrieved. This kind of system is applied in a real case, and it allows teachers and managers of the MOOC platform to get enhanced information and insights about users interaction with contents and MOOC tools, as well as some metrics impossible to retrieve or calculate manually in this kind of eLearning platforms with high amounts of users.

In order to demonstrate the utility of this kind of software architectures, showing also the possibilities and new metrics that could be gathered using it, authors show some data gathered from the iMOOC platform and social networks, showing in these data how the application of this software architecture can help to measure elements difficult to estimate because the vast amount of users enrolled in MOOC courses or because they implicate the usage of external tools like the social networks that teachers can not track without this kind of tools and systems.

8. ACKNOWLEDGMENTS

Authors would like to thank to the GRIAL and LITI research groups by their support in this research, also they would like to thank to the Technical University of Madrid by its financial support throughout the "Diseño y desarrollo de MOOC universitarios" (Ref. PT1415-05000) Project and to the Vice Chancellor for Academic Policy at the University of Salamanca for financing the project under the Teaching Innovation program ID2014/0281 that allowed the authors to develop the software architecture described in this research work.

9. REFERENCES

[1] Alario-Hoyos, C., Pérez-Sanagustín, M., Delgado-Kloos, C., Parada G, H., Muñoz-Organero, M., and Rodríguez-De-Las-Heras, A., 2013. Analysing the Impact of Built-In and External Social Tools in a MOOC on Educational Technologies. In Scaling up Learning for Sustained Impact, D. Hernández-Leo, T. Ley, R. Klamma and A. Harrer Eds. Springer Berlin Heidelberg, 5-18. DOI= http://dx.doi.org/http://dx.doi.org/10.1007/978-3-642-40814-4_2.

[2] Bjornavold, J. and European Centre for the Development of Vocational Training, 2008. Validation of non-formal and informal learning in Europe: a snapshot 2007 Office for Official Publications of the European Communities, Luxembourg.

[3] Borrás Gené, O., Martínez Núñez, M., and Fidalgo Blanco, Á., 2014. Gamification in MOOC: challenges, opportunities and proposals for advancing MOOC model. In Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality, F.J. García-Peñalvo Ed. ACM, New York, USA, 215-220. DOI= http://dx.doi.org/10.1145/2669711.2669902.

[4] Borrás-Gené, O., 2013, Aplicación de las redes sociales a la enseñanza: Comunidades virtuales. . Available https://www.miriadax.net/web/aplicacion-rrss-ensenanza. Retrieved 2/04/2013

[5] Cabero, J., 2006. Las TIC y las inteligencias múltiples. Infobit. Revista para la difusión y uso educativo de las TIC 13, 75.

[6] Coursera, 2015, Coursera MOOCs Homepage. Available http://coursera.org/. Retrieved 3/05/2015

[7] Cruz-Benito, J., García-Peñalvo, F.J., and Therón, R., 2014. Defining generic data collectors for Learning Analytics: Facing up the heterogeneous data from heterogeneous environments. In Proceedings of the International Conference on Advanced Learning Technologies (ICALT) 2014 (Athens, Greece2014), IEEE. DOI= http://dx.doi.org/10.1109/ICALT.2014.108.

[8] Cruz-Benito, J., Garcia-Peñalvo, F.J., Theron, R., Maderuelo, C., Perez-Blanco, J.S., Zazo, H., and Martin-Suarez, A., 2014. Using software architectures to retrieve interaction information in eLearning environments. In Computers in Education (SIIE), 2014 International Symposium on, 117-120. DOI= http://dx.doi.org/10.1109/SIIE.2014.7017715.

[9] Cruz-Benito, J., Therón, R., García Peñalvo, F.J., Maderuelo, C., Pérez-Blanco, J.S., Zazo, H., and Martin-Suarez, A., 2014. Monitoring and feedback of Learning Processes in Virtual Worlds through analytics architectures: A real case. In Sistemas y Tecnologías de Información. Actas de la 9ª Conferencia Ibérica de Sistemas y Tecnologías de Información, Á. Rocha, D. Fonseca, E. Redondo, L.P. Reis and M.P. Cota Eds. AISTI (Asociación Ibérica de Sistemas y Tecnologías de Información), Barcelona, España, June, 18-21, 2014, 1126-1131.

[10] Cruz-Benito, J., Therón, R., García-Peñalvo, F.J., and Pizarro Lucas, E., 2015. Discovering usage behaviors and engagement in an Educational Virtual World. Computers in Human Behavior 47, 18-25. DOI= http://dx.doi.org/10.1016/j.chb.2014.11.028.

[11] Csikszentmihalyi, M. and Hermason, K., 1995. Intrinsic Motivation in Museums-What Makes Visitors Want to Learn. Museum news 74, 3, 34-37& 59-62.

[12] Django Software Foundation, 2015, Django Web Framework. Available https://www.djangoproject.com/. Retrieved 15/03/2015

[13] Edx, 2015, Edx MOOCs Homepage. Available https://www.edx.org/. Retrieved 3/05/2015

[14] Efron, M., 2010. Hashtag retrieval in a microblogging environment. In Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval ACM, 787-788.

[15] Eshach, H., 2007. Bridging in-school and out-of-school learning: Formal, non-formal, and informal education. Journal of science education and technology 16, 2, 171-190.

[16] Faure, E., 1972. Learning to be: The world of education today and tomorrow. Unesco.

[17] Fidalgo Blanco, Á., García-Peñalvo, F.J., and Sein-Echaluce Lacleta, M.L., 2013. A methodology proposal for developing adaptive cMOOC. In Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality, F.J. García-Peñalvo Ed. ACM, New York, USA, 553-558.

[18] Fidalgo Blanco, Á., Sein-Echaluce Lacleta, M.L., and García-Peñalvo, F.J., 2013. MOOC cooperativo. Una integración entre cMOOC y xMOOC. Cooperative MOOC. An integration between cMOOC and xMOOC. In II Congreso Internacional sobre Aprendizaje, Innovación y Competitividad, CINAIC 2013, Á. Fidalgo Blanco and M.L. Sein-Echaluce Lacleta Eds. Fundación General de la Universidad Politécnica de Madrid, Madrid, 481-486.

[19] Fielding, R.T., 2000. Architectural Styles and the Design of Network-based Software Architectures. In Irvine University of California.

[20] García-Peñalvo, F.J., Cruz-Benito, J., Borrás-Gené, O., and Fidalgo Blanco, Á., 2015. Evolution of the Conversation and Knowledge Acquisition in Social Networks related to a MOOC Course. In Proceedings of the HCI International (Los Angeles, CA, USA, In press 2015), Springer.

[21] García-Peñalvo, F.J., Cruz-Benito, J., Maderuelo, C., Pérez-Blanco, J.S., and Martín-Suárez, A., 2014. Usalpharma: A Cloud-Based Architecture to Support Quality Assurance Training Processes in Health Area Using Virtual Worlds. The Scientific World Journal 2014.

[22] Google, 2015, Google+ Platform APIs. Available https://developers.google.com/+/api/. Retrieved 1/03/2015

[23] Huang, J., Thornton, K.M., and Efthimiadis, E.N., 2010. Conversational tagging in twitter. In Proceedings of the 21st ACM conference on Hypertext and hypermedia ACM, 173-178.

[24] Miriadax, 2015, MiriadaX MOOCs Homepage. Available http://miriadax.net/. Retrieved 3/05/2015

[25] Mongodb, 2015, Website. Available https://www.mongodb.org/. Retrieved

[26] Moodle, 2014, Web service API functions. Available https://docs.moodle.org/dev/Web_service_API_functions. Retrieved 1/03/2015

[27] Moodle, 2015, Website. Available http://moodle.org. Retrieved 3/04/2015

[28] Núñez, M.M., Gené, O.B., and Blanco, Á.F., 2014. Social community in MOOCs: practical implications and outcomes. In Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality ACM, 147-154.

[29] Project, T., 2015, Tweepy on Github. Available https://github.com/tweepy. Retrieved 1/03/2015

[30] Requests, 2015, Requests: HTTP for Humans. Available http://docs.python-requests.org/. Retrieved 1/02/2015

[31] Twitter Inc, 2015, REST APIs. Available https://dev.twitter.com/rest/public. Retrieved 1/03/2015

[32] Universidad Politécnica De Madrid, Universidad De Zaragoza, and Universidad De Salamanca, 2015, iMOOC. Available http://gridlab.upm.es/imooc/. Retrieved 1/03/2015.

6.7 Appendix G. Learning communities in social networks and their relationship with the MOOCs

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RITA.2017.2655218, IEEE Revista Iberoamericana de Technologias del Aprendizaje

Learning communities in social networks and their relationship with the MOOCs

Juan Cruz-Benito, Oriol Borrás-Gené, Francisco J. García-Peñalvo, Ángel Fidalgo Blanco, Roberto Therón

Abstract— **This article discusses how MOOC users learn and participate in cooperative environments that promote learning communities within external hypermedia environments such as the social networks. In order to develop this study, researchers analyzed the interaction of users enrolled in a course developed under the iMOOC platform, which is based on concepts like connectivism, collaborative learning, gamification, or adaptivity, among others. Specifically, this study deals with obtaining information about the conversations that take place in external learning communities within social networks like Google+ and Twitter in parallel with the iMOOC platform itself. Through this information is possible to establish the learning types that experience users (non-formal and informal learning usually) and providing an estimation of how users interact with content tagged in social network, and how they use these tagging facilities to continue or create new conversations that allow them to expand or strengthen their learning process developed in the MOOC. To complete the knowledge extracted from these tags and to understand how users interact with this way of metadata declaration, the study is complemented by a questionnaire that collected how users utilize and understand of these tags based on the main usages and the age of the users of the MOOC.**

Index Terms— **Interaction, HCI, Analysis, E-learning, MOOCs, Collaborative learning, Social networks, Knowledge society, Software Architectures**

I. Introduction

The emergence of the Internet and the eLearning concept have modified the way humans learn and interact with knowledge [1-5]. Specifically, this change has undergone a marked process of acceleration with the emergence of new theories, methodologies, tools and systems. These emergent items, designed and implemented to take more and better advantage of the online medium, facilitate the acquisition of knowledge and learning by students, without minding age, sex or other personal conditions [6, 7]. The evolution of technological ecosystems [8] intended to support eLearning has given rise to environments capable of adapting to the users [9-11] providing gamification techniques [12-14] to the process of acquiring knowledge, offering collaborative learning tools [15], analyzing their learning and acting accordingly to reinforce it [16, 17], or providing the ways to multiply the number of Students who can study a subject simultaneously on a virtual platform [18]. Especially in fashion are these last learning environments, better known as MOOC (Massive Online Open Courses) [19]. These learning environments make available to many users (thousands in many cases) virtual classes as knowledge containers that provide open learning resources for all users enrolled in the course, as well as methods and systems to reinforce that knowledge acquisition from different perspectives [20-22].

On the other hand, in addition to the revolution that has led eLearning in the field of learning, it is necessary to emphasize that the boom in systems and methodologies such as those used to support the MOOCs is due in large part to the fact that learning is not currently conceived from the classical formal point of view. According to the literature [23-30], at present, three types of learning can be distinguished from the concept of learning:

- Formal learning is “that which occurs in organized and structured environments (in an educational institution or in a work environment) and is explicitly designed as learning in terms of objectives, time and resources. Typically leads to processes of validation and certification on the knowledge acquired”.
- Non-formal learning is “one that is embedded in planned activities without being explicitly posed as learning. Nonetheless, non-formal learning can occasionally be validated and lead to certifications”.
- Informal learning is “one that results from daily activities related to work, family, or leisure. It is an unorganized and unstructured learning in terms of objectives, timing or support of typical structures related to learning”.

The combination of these types of learning (especially non-formal and informal) with systems and methodologies purely related to eLearning have opened, and open today, new perspectives on the formation of the individual of the 21st

Juan Cruz-Benito and Francisco J. García-Peñalvo are with the GRIAL Research Group, Computer Science Department, Research Institute for Educational Sciences, University of Salamanca, Salamanca 37008, Spain (e-mails: juancb@usal.es and fgarcia@usal.es).

Oriol Borrás-Gené Is a member of the Tele-Education Cabinet (GATE) of the Universidad Politécnica de Madrid (e-mail: oriol.borras@upm.es)

Ángel Fidalgo Blanco Is Director of the Laboratory of Innovation in Information Technology of the Universidad Politécnica de Madrid, España (e-mail: afidalgo@dmami.upm.es)

Roberto Theron is with the GRIAL Research Group, Computer Science Department, and Vis-USAL Research Group, University of Salamanca, Salamanca 37008, Spain (email: theron@usal.es)

DOI (Digital Object Identifier) Pending

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RITA.2017.2655218, IEEE Revista Iberoamericana de Technologias del Aprendizaje

century from a continuous point of view in time, multivariate in terms of resources that can consume and ubiquitous in terms of availability. At the present time, where information overflows the individual, having the means and vehicles to drive true and valuable knowledge to the people can be one of the cornerstones in building a more wise and advanced society.

Among these information vehicles, a number of hypermedia resources have appeared over the years. Among them the most important are web sites, forums, blogs, and from several years, the social networks [31]. These social networks support true learning communities [32] where conversations are developed, content is shared openly, relationships are established between users (in a horizontal way in many cases), appear interaction between people and digital entities, etc. In other words, social networks constitute a real *Petri dish* where users form digital societies by developing communication structures, consumption patterns (in this case information) and user networks in a broad way. Is in these social networks, in these digital societies, where some of the aspects mentioned previously can come together: eLearning in a non-formal or informal context, and also formal [33]. As discussed by various authors [34-36], informal conversations and content present in social networks is currently one of the most successful ways to acquire extra knowledge and improve the learning experience in online courses. On the other hand, certain conversations and interactions that take place in social networks can be a consequence of the realization by social network users of some kind of learning activity (online or offline). [37]. In this sense, it is possible to cite the Connectivism theory [38, 39], which promulgates that the learning process is enriched by the connection of students, teachers and online resources, and to which it can be added that social networks are a perfect means to improve This connection [40, 41], thus favoring the emergence of true communities of connected learning and practice [42].

The present manuscript deals exactly with these last questions. The study of the informal conversations is considered, and the interaction of the students of a course of a MOOC with contents of the MOOC itself, in a non-formal and informal way [43-48]. The MOOC platform that has been used to carry out the study is iMOOC: a system developed by the Polytechnic University of Madrid, the University of Zaragoza and the University Salamanca. iMOOC is based on non-formal and informal learning and has characteristics of adaptability, gamification, or collaborative learning among others. A more complete description is given in Section 2 (Materials and Methods).

In this way, the objectives of this article could be summarized as follows:

1. Study user conversations and interactions with MOOC content or related on social networks like Google+ or Twitter and how users use tagging resources on such networks (hashtags) [49].
2. To determine if there are patterns or coincidences between the use of social networks and the development of the MOOC by the users.
3. To evaluate if it is possible to use the performance of users in MOOCs and social networks to establish parallels between both and to determine the types of learning that are given in these environments in order to be able to take advantage of that knowledge through the feedback of the educational environment MOOC.

These objectives, as well as the rest of the content and concepts needed to work with them, are discussed in the following sections: Materials and Methods, Results, Discussion and Conclusions.

II. Materials and Methods

A. Materials

1) iMOOC

The iMOOC or intelligent-MOOC platform [50, 51] includes, among other actions, the creation of a MOOC platform based on adaptive learning and information [26, 52-54]. To achieve this goal, the project uses of the eLearning Moodle platform (https://docs.moodle.org/dev/Core_APIs), specifically version 2.6.5, taking advantage of its great versatility.

This adapted learning is possible thanks to the use of the different tools offered by the platform. These tools include conditionals, groups and user groups that allow creating and then choosing different groups by associating them with the different course resources. The use of these strategies and tools can lead to different educational itineraries depending on the type of profile of the user, path chosen according to the theme or the progress of the student within the course and level of knowledge.

2) iMOOC Course "Social Networks and Teaching"

On January 12, 2015, a first demo of the iMOOC platform was launched with the course "Social Networks and Teaching", a special version of the course "Application of social networks to teaching" previously developed in the platform MiriadaX (https://miriadax.net/web/aplicacion-rrss-ensenanza).

This course is based on the cooperative model defined by Fidalgo *et. al.* [55] which contains characteristics of the two more standardized types of MOOCs such as xMOOC, with a more behavioral and similar approach to traditional online courses compared to cMOOCs whose approach is more connectivist [56, 57] based on the social networks. In order to explain this cooperative model, the course can be divided into a series of layers, starting with the "technological" layer. This layer includes, on the one hand, the MOOC platform that hosts the course and on the other, the social platforms where the interactions between participants and the content generation by these too. This layer is followed by the "formative strategy" associated with the instructional design of the course and finally the "cooperative" layer that represents the most connectivistic part of the course, gathering the results and the content generated from the cooperation between the teaching team and the participants in the course and integrating it with this. It is necessary to add to the defined cooperative model a

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RITA.2017.2655218, IEEE Revista Iberoamericana de Technologias del Aprendizaje

fourth layer to explain the gcMOOC cooperative-gamified model on which the course is based, this layer is called gamification [58], which interacts with the other layers, fostering the motivation of the participants in the course [58].

As for the content of the course, it aims to introduce the student to the social web for a month, identifying the phenomenon of social networks within it and the opportunities they offer within the field of education, more specifically, in the generation of Virtual learning communities. On the other hand, the use of the most widespread social networks such as Facebook, Twitter and Google+ is explored in order for the student to develop the digital skills necessary to deal with this type of technology, while offering a series of guidelines for application in the classroom. Finally, it gives an overview of a total of 13 other social networks also suitable for this teaching use, as well as tools for a more optimal management.

Based on the characteristics offered by the iMOOC platform, students are offered five different itineraries that they can choose before starting the course, adapting it to their needs:

- Complete course for teachers (with two additional lessons focused on the use of Twitter and Facebook as an educational tool).
- Complete course for non-teachers (without teaching lessons).
- Twitter course, formed only by the module of said social network.
- Facebook course, a single module on this network.
- Special itinerary, focused on those students who participated in any of the previous editions of the course, with an additional module focused on the practical use of the learning community. This itinerary also offered the possibility of repeating the complete course, allowing its students to access all the contents.

As general data of the course follow-up and some metrics about the itineraries and completion of the course, the following summary is offered to the reader (Figure 1, Figure 2, Figure 3). In the specific case of Figure 3, those students who chose the special route to obtain the certificate were excluded from the metrics, since they had only to present the certificate obtained in either of the previous editions of the course in MiriadaX, to visualize the contents of its special module, based on an implementation of learning communities from two different points of view and participate in the forum. Even so, the students of the special itinerary were offered the possibility of completing the full MOOC. Of the 188 who made this itinerary, 107 repeated the course and of them, 88 passed it (82.2%).

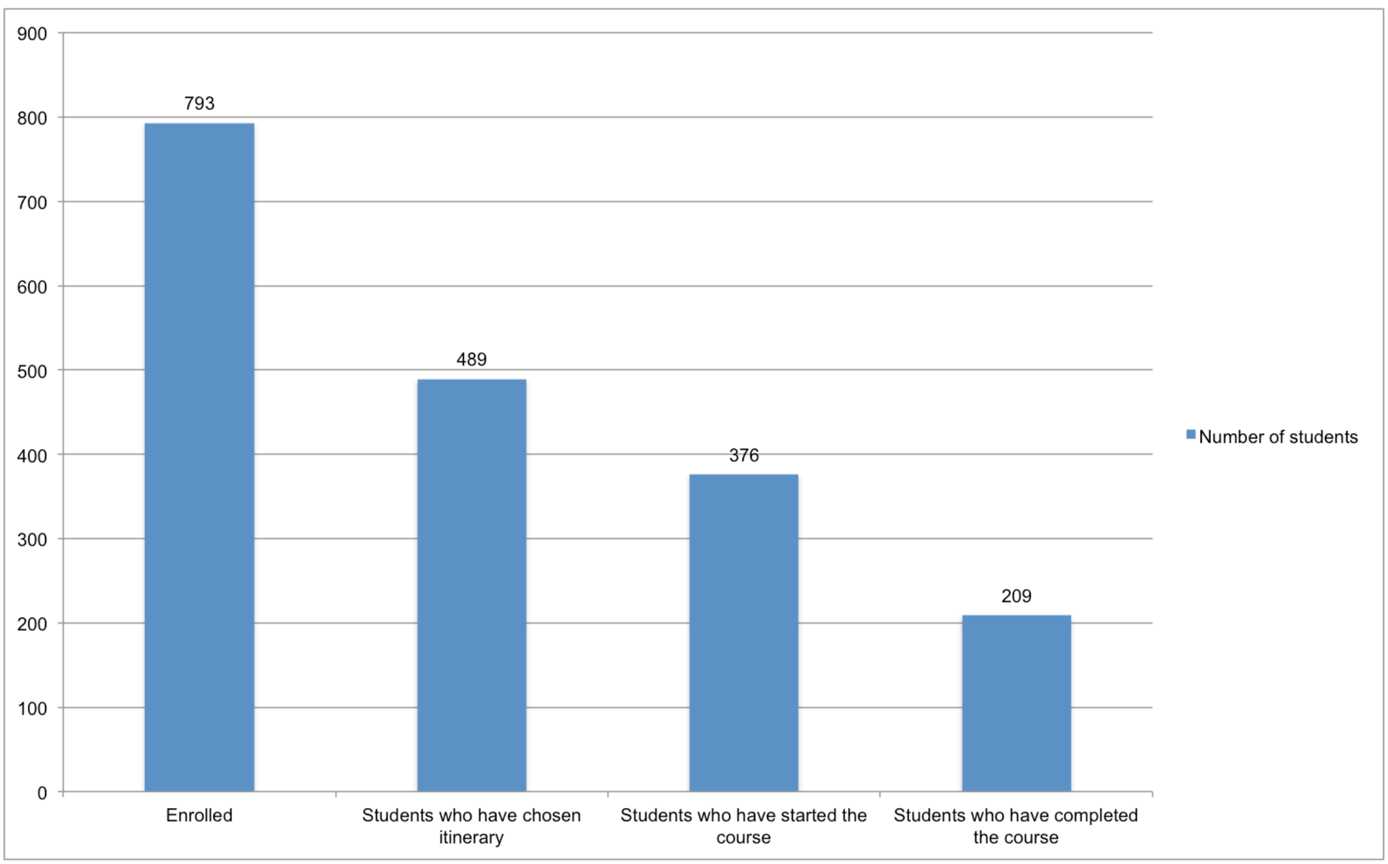


Figure 1. Distribution of students regarding enrollment, choice of itinerary, initiation and completion of the MOOC course.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RITA.2017.2655218, IEEE Revista Iberoamericana de Technologias del Aprendizaje

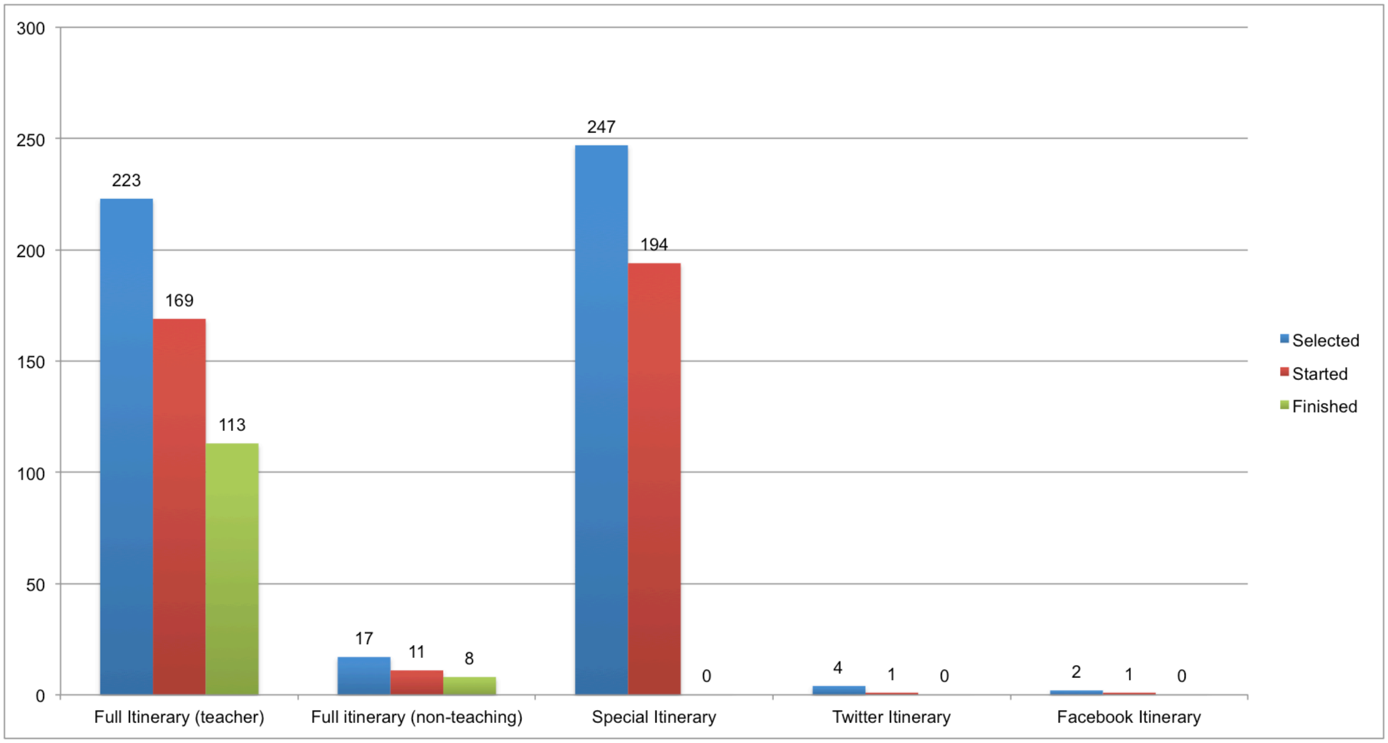


Figure 2. Distribution of students regarding the learning itineraries available in the course

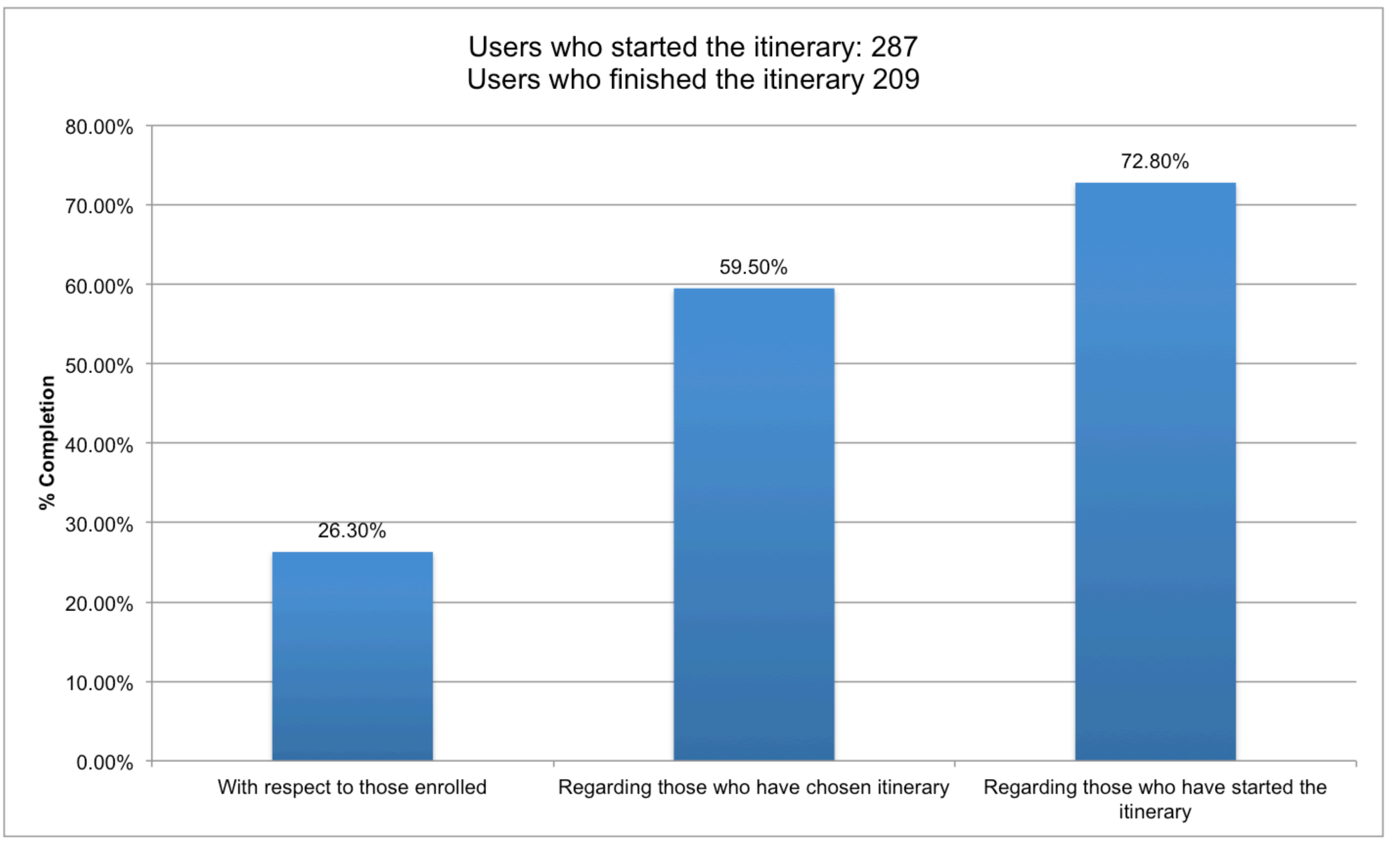


Figure 3. Statistics on the initiation of completion of the MOOC course without taking into account the special itinerary, because this itinerary did not require the same effort as the rest.

3) Social networks: Google+ and Twitter

In the case of a course that deals with "social networks in teaching", and given the connectivist approach of the MOOC and the course, the social networks have played an important role in the learning process associated to the course. These networks have been used from two different perspectives. The first is the use of social networks as test environments and as case studies to obtain a practical understanding of the concepts

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RITA.2017.2655218, IEEE Revista Iberoamericana de Technologias del Aprendizaje

theoretically shown in the course. The second perspective is its use as a platform to continue and extend the learning process of the iMOOC course from both a non-formal and informal perspective. To accomplish this second perspective, teachers proposed suitably labeled conversation topics (through hashtags), opening up ways of discussion and acquisition of knowledge from a non-formal point of view, encouraging the emergence of conversations and informal learning. The communication channels were proposed for the students themselves in an environment other than the MOOC, as well as for discussion with other social network users who participate in the conversation without being enrolled in the course [59] as shown in Figure 4.

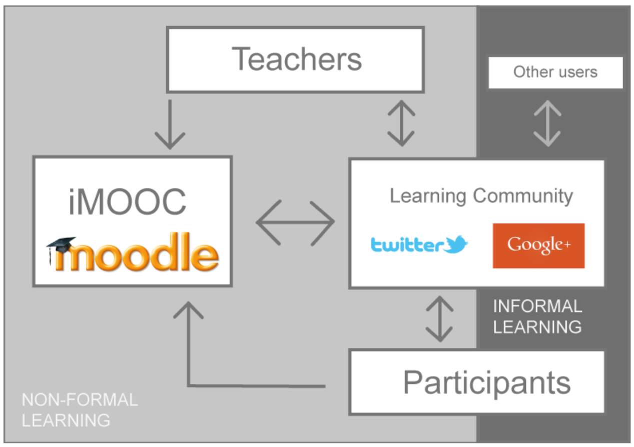


Figure 4. Schema on the interaction between users, iMOOC and social networks segmenting learning according to the type that occurs in each phase.

For this task, this MOOC has used social networks Google+ and Twitter. In Google+ this course has been associated, since its inception in the MiriadaX platform, to a community with more than 5000 users (https://plus.google.com/u/0/communities/110109268598525419924). In this community, users share resources and collaborate among them in the learning associated with each edition of the course. Regarding to Twitter, it has been used as an alternative social network to raise discussions and conversations between users, due to its current popularity and the facilities it offers regarding to tag, track and information retrieval from users' conversations [60-62]. The following sections (Methods, Results, Discussion and Conclusions) show how this process of tracking, retrieving and exploring the data has been carried out to achieve the objectives set for both Twitter and Google+.

B. *Methods*

In order to obtain information about the interaction of iMOOC users with the contents of the course in social networks, it is necessary to establish adequate information retrieval channels [63]. Specifically in this case about the relationship between iMOOC user profiles and different social networks, as well such as retrieving the information they share and the tagging they perform on those contents, in order to be able to carry out the analyzes that are presented as study objectives.

As a summary of the methods used, can be highlighted:

- In order to largely avoid manual retrieval of each user's data, web services have been used (REST APIs http://www.ics.uci.edu/~fielding/pubs/dissertation/rest_arch_style.htm) that offers the Moodle platform on which the iMOOC is based. Using REST API, user registries and their profiles have been accessed, allowing authors to filter those who have registered their profiles in social networks (condition proposed by the MOOC teachers to help in the evaluation of students) [62].
- To extract information from Twitter, the automatic recovery of tweets (through its REST API https://dev.twitter.com/rest/public) has been combined with the manual recovery of some specific metrics.
- As for extracting information from the social network Google+, due to the lack of APIs to retrieve information from user communities (https://developers.google.com/+/web/api/rest/), it has been necessary Develop a tool called GILCA (Google Analytics Informal Learning Communities). This application collects data from Google+ communities through the email notifications sent by the social network (including information on publications, comments, hashtags, etc.) [62].
- To further understand how users use the tagging functions in the social networks used in this article, they conducted a questionnaire that addressed basic questions about the use of hashtags in social networks and in course related activities.

As for the analysis of the data obtained, this was done with spreadsheet tools.

III. Results

A. *Results of the iMOOC questionnaire*

As mentioned, the questionnaire filled out by course users contained questions related to the use of labels on social networks [64]. 212 users, 26.73% of the total enrolled in the course, completed this questionnaire. In the following figures (Figure 5, Figure 6, Figure 7, Figure 8, Figure 9, Figure 10) can be seen the data retrieved through these questionnaires and the segmentation of responses according to the age of the users.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RITA.2017.2655218, IEEE Revista Iberoamericana de Technologias del Aprendizaje

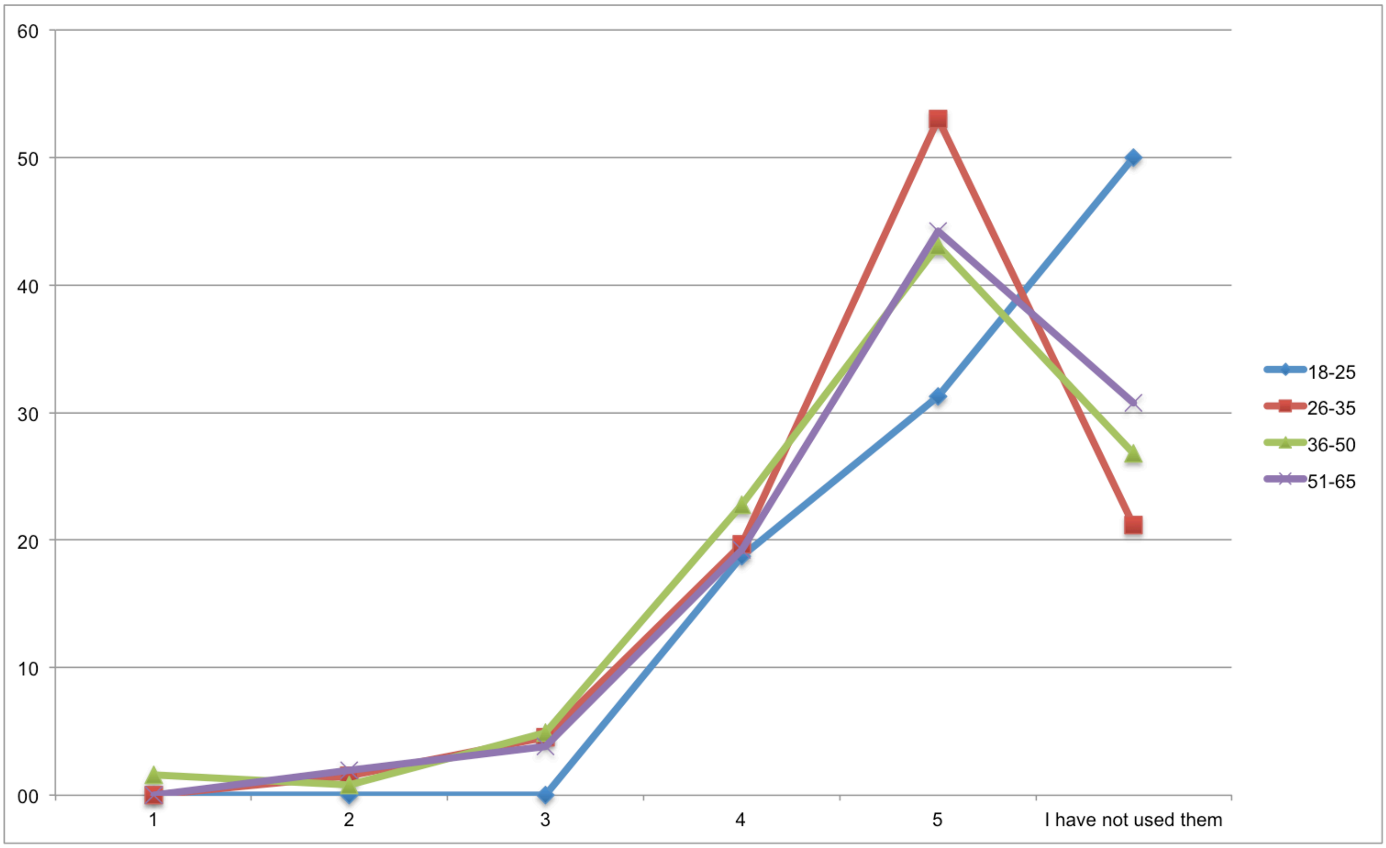


Figure 5. Results of question 1 of the iMOOC questionnaire: Evaluates the utility of using hashtags (Likert scale 1-5 plus null value). Age-targeted responses of users

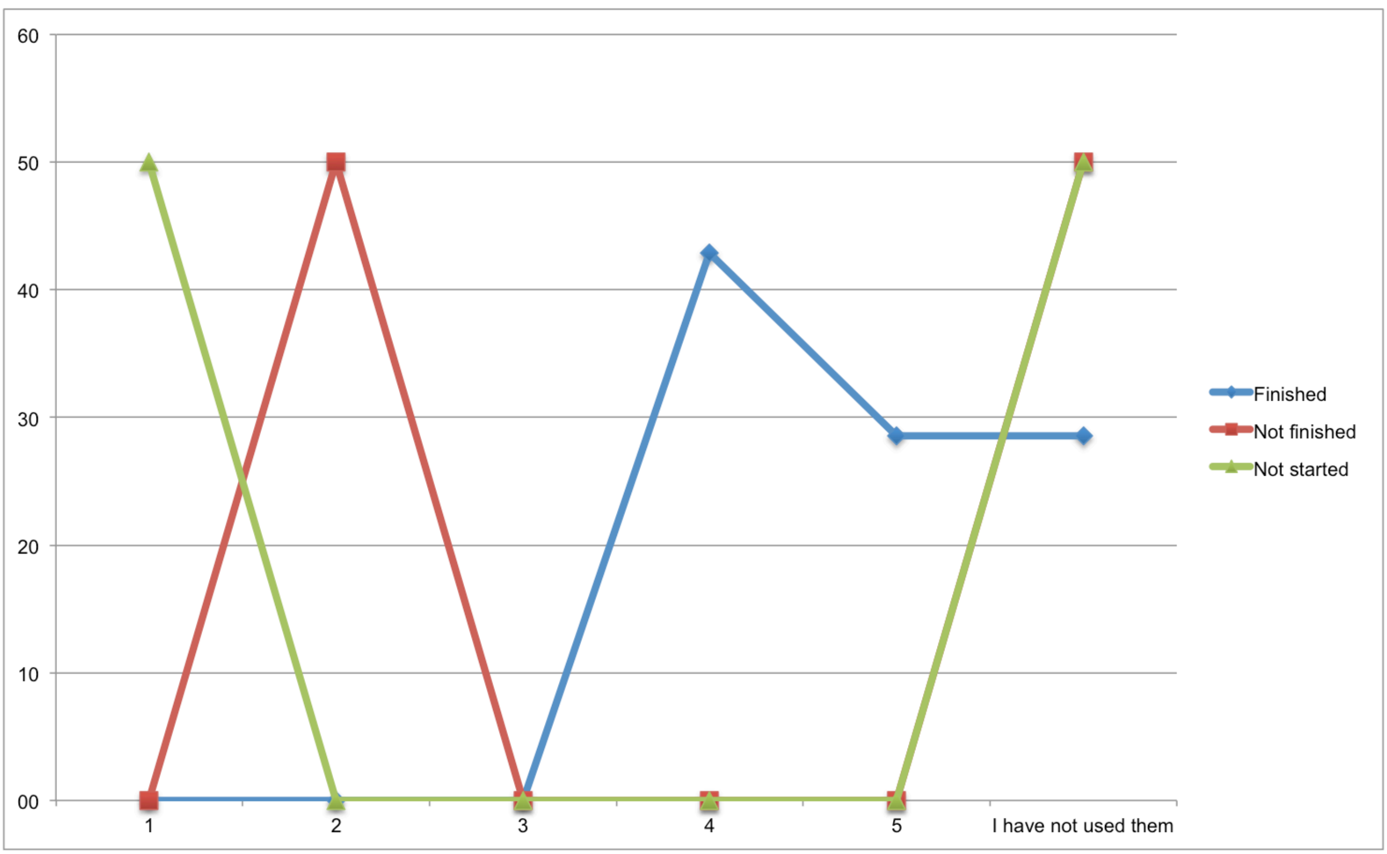


Figure 6. Results of question 1 of the iMOOC questionnaire: Evaluates the utility of using hashtags (Likert scale 1-5 plus null value). Responses segmented by degree of completion of the MOOC by the users.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RITA.2017.2655218, IEEE Revista Iberoamericana de Technologias del Aprendizaje

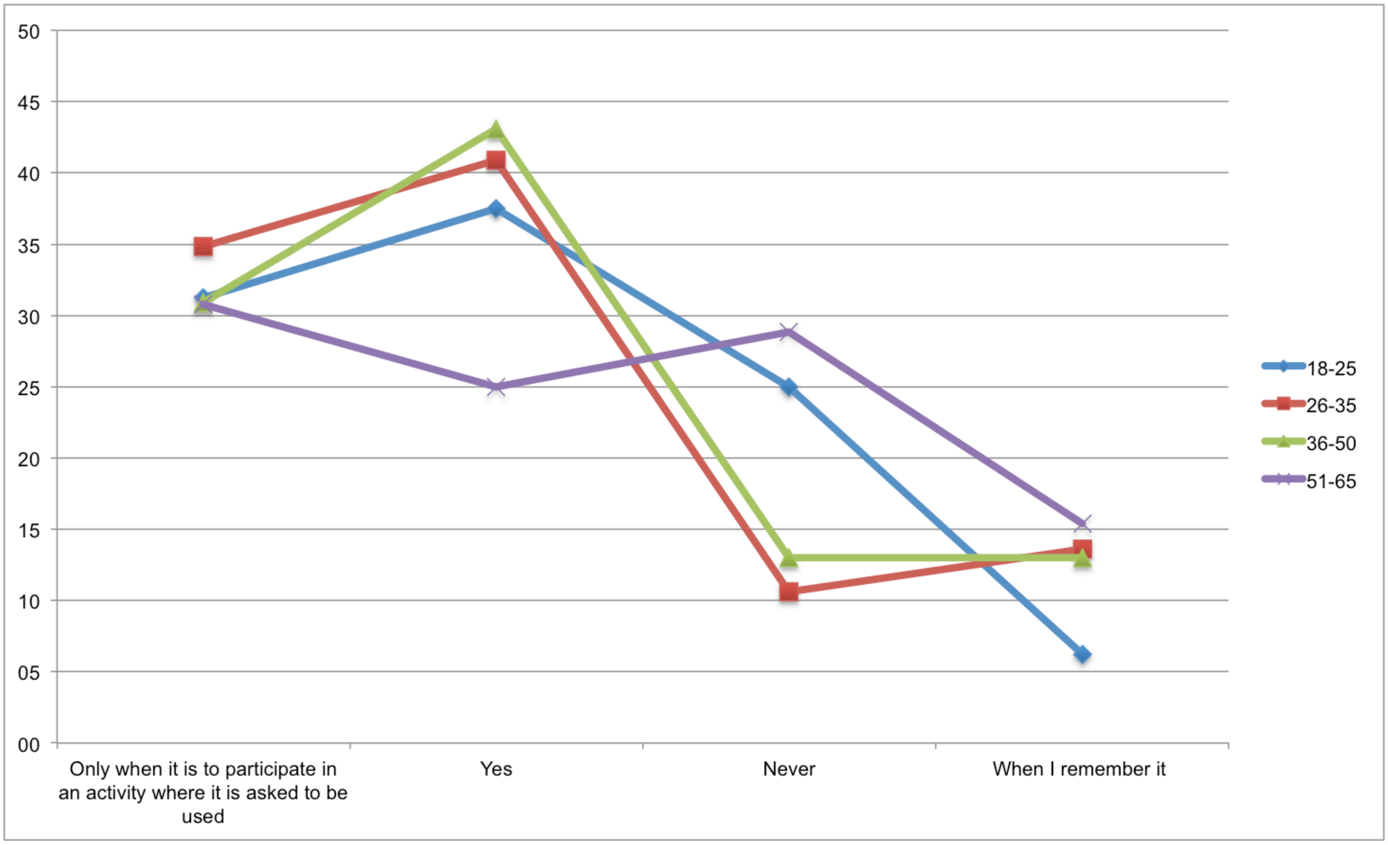


Figure 7. Results from question 2 of the iMOOC questionnaire: Do you often use hashtags in your publications? Answers segmented by age groups of users.

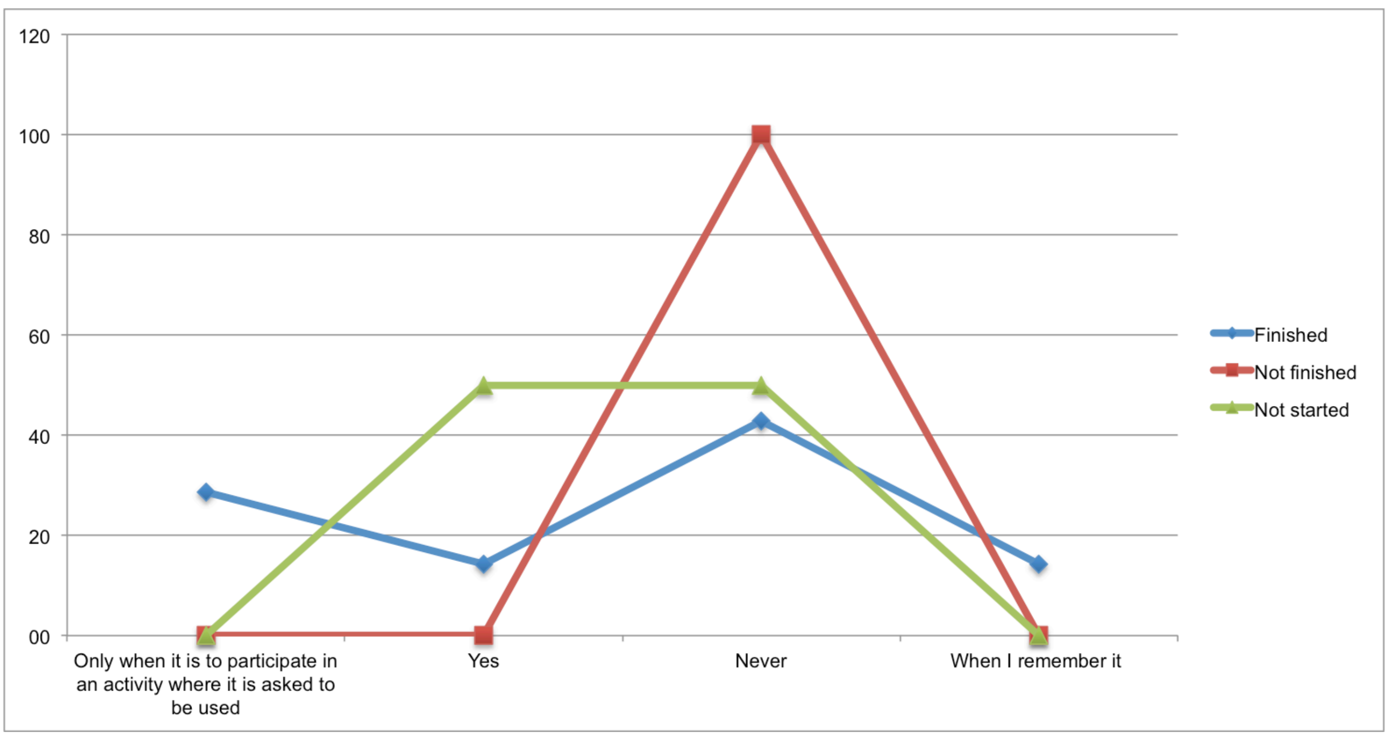


Figure 8. Results from question 2 of the iMOOC questionnaire: Do you often use hashtags in your publications? Responses segmented by degree of completion of the MOOC by the users.

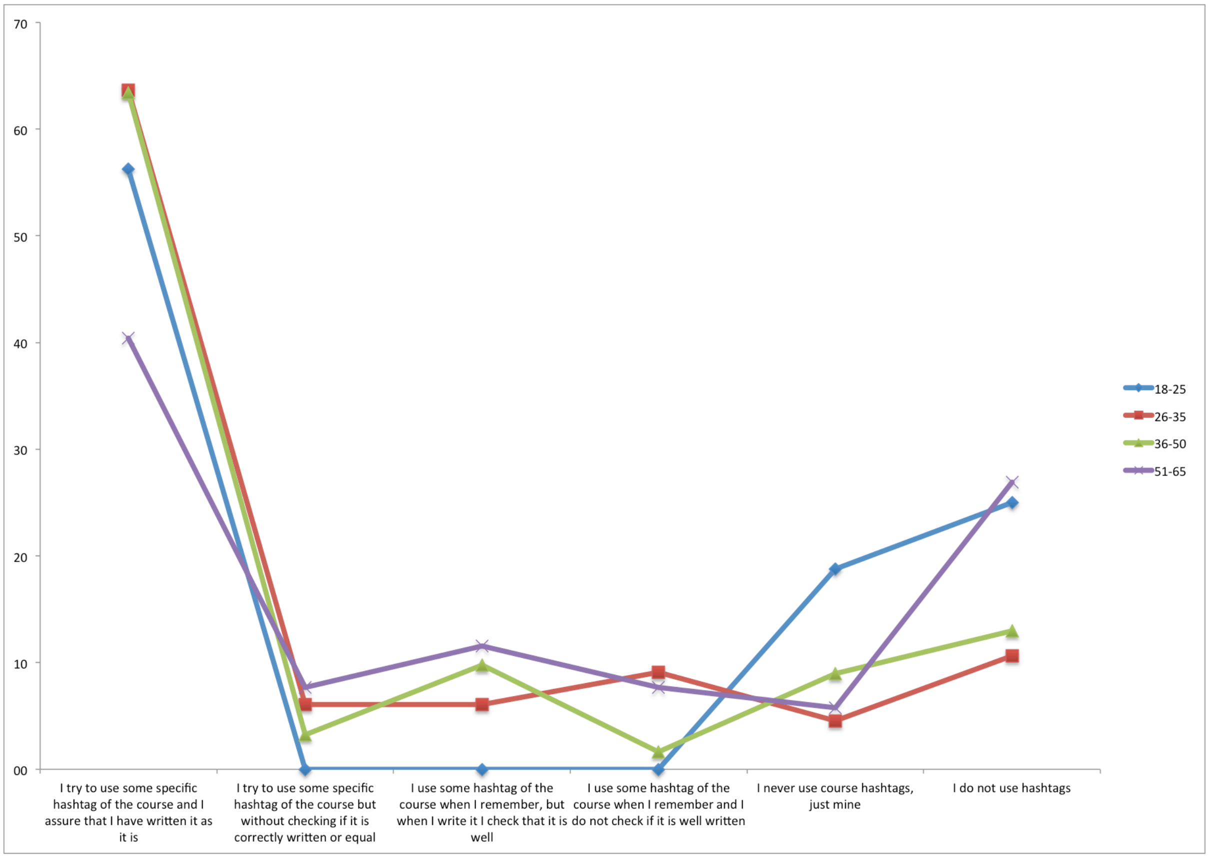


Figure 9. Results from question 3 of the iMOOC questionnaire: "When publishing and using a hashtag ...". Answers segmented by age groups of users.

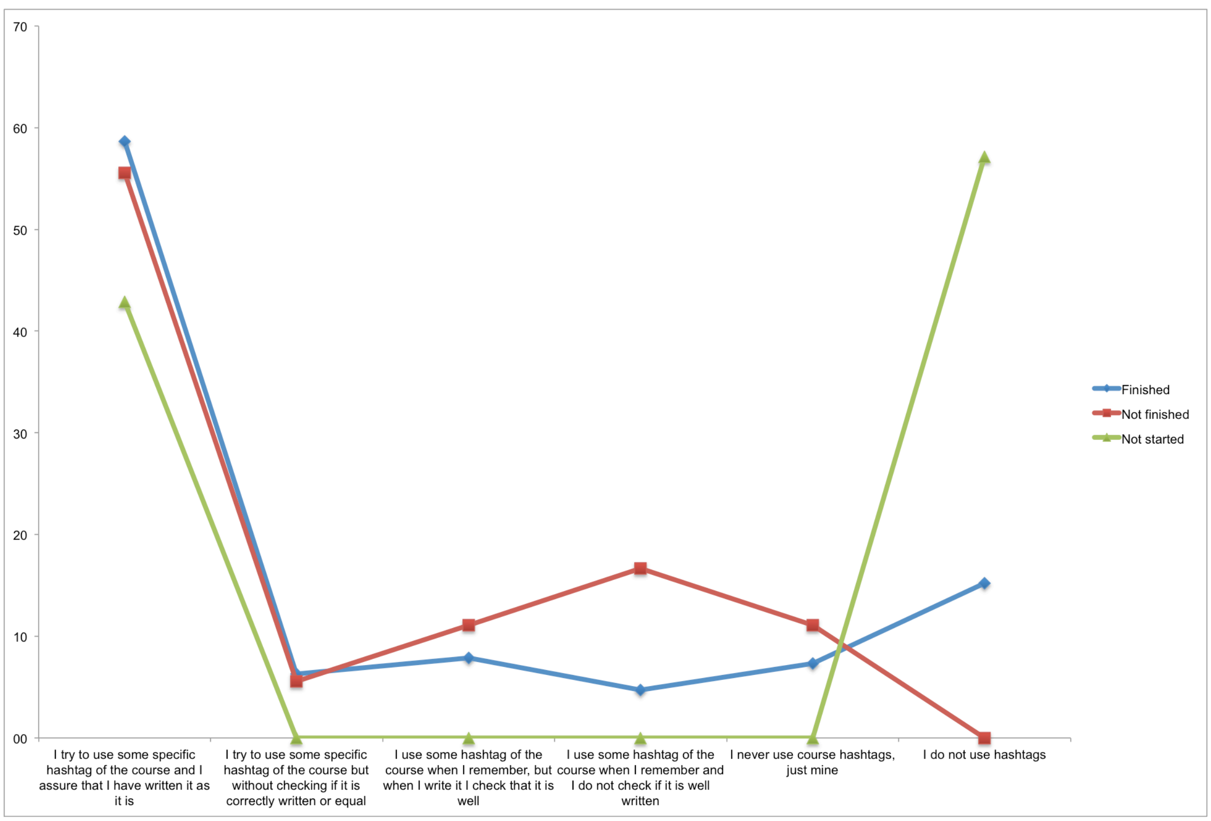


Figure 10. Results from question 3 of the iMOOC questionnaire: "When publishing and using a hashtag ...". Responses segmented by degree of completion of the MOOC by the users.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RITA.2017.2655218, IEEE Revista Iberoamericana de Technologias del Aprendizaje

B. *Results obtained in social networks*

In order to obtain results from the use of social networks, researchers filtered what users had indicated their Google+ profile or Twitter in their iMOOC profile. This filter allowed to retrieve which of them had been published in social networks following the official hashtags of the course (Figure 11). Once this check was carried out, the researchers evaluated the number of publications they had made and the proportion of users published on Google+ or Facebook. Finally was evaluated if these users with social interactions and social profiles had approved the course or not. This evaluation resulted in:

- The students enrolled in the course have made a total of 263 publications in the Google+ community, also commenting on other peer publications.
- The students enrolled in the course have published a total of 131 tweets following official and unofficial hashtags on Twitter.
- Of the users who have made a publication in Google+ (191 users), have passed the course 57 (29.84%).
- Of the users who have published a tweet on Twitter (76), have passed 42 (55.26%) the course.
- Of the 191 users who indicated their Google+ profile on iMOOC, 83 users (43.5%) have passed the course.
- Of the 265 users who have indicated their Twitter profile on iMOOC, 105 users (39.62%) have passed the course

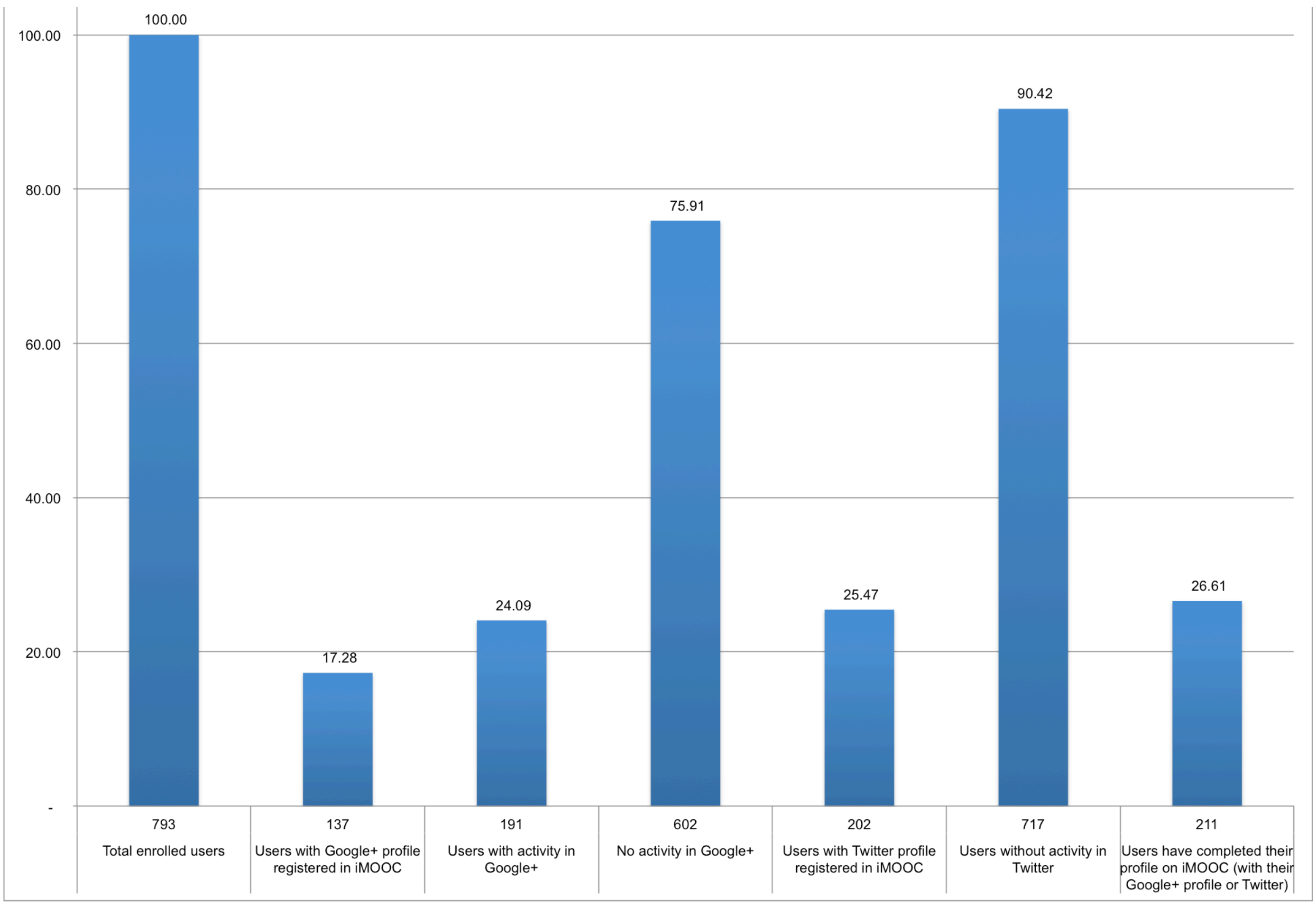


Figure 11. Distribution of iMOOC users regarding their use in social networks

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RITA.2017.2655218, IEEE Revista Iberoamericana de Technologias del Aprendizaje

Table 1. Distribution of interactions on Google+ and Twitter by type of content and learning

Google +						
Publication						
Category	**Type**	**Amount**	**+1s**	**Comments**	***Reshares***	**Type of learning**
Debates	Proposed throughout the course	1	83	17	14	Non formal (proposed by teachers)
	Use of social networks	4				Informal (proposed by students)
	About learning	3				
	About digital identity	2				
	About digital identity	1				
	About Facebook	1				
	About *badges*	1				
Total posts in the discussion category = 11						
Activities and exercises	Examples of social networks	31	309	41	20	Non formal (Proposed by teachers)
	Exercises on bad practices in social networks	25				
	Exercises about Facebook	28				
	About influence (Klout)	22				
	Uses of Twitter in teaching	3				
	Others	2				
Total publications on activities and exercises = 111						
Resources	150		552	66	93	Informal (proposed by the students)
Twitter						
Publication						
Type	**Hashtag**	**Tweets**	**Responses**	**Retweets**	**Favorites**	**Type of Learning**
General	#RSEMOOC	9	2	5	5	Non formal (proposed by teachers)
	#RSEHANGOUT	19	4	16	15	
	#Modulo1RSE	1	0	1	1	
	#Modulo2RSE	1	0	1	1	
	#Modulo3RSE	1	0	1	1	
Activities and exercises	#RSEEjemplosRRSS	4	1	0	0	
	#RSEMalasPracticas	5	0	1	2	
	#RSEmiKlout	8	1	5	6	
	#RSEMoodleTwitter	59	9	9	11	
	#ActividadesRSE	1	0	3	3	
Total tweets = 107						

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RITA.2017.2655218, IEEE Revista Iberoamericana de Technologias del Aprendizaje

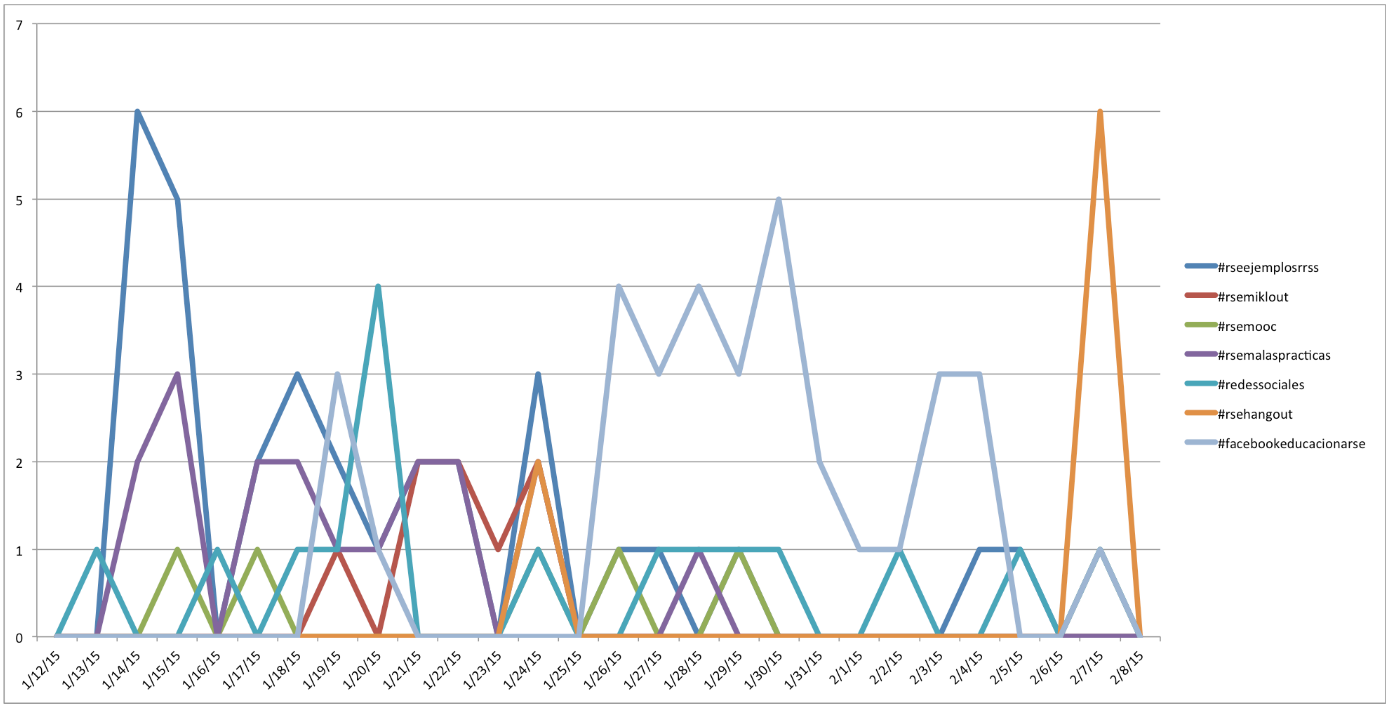


Figure 12. Evolution of the use of non-formal hashtags on Twitter throughout the course.

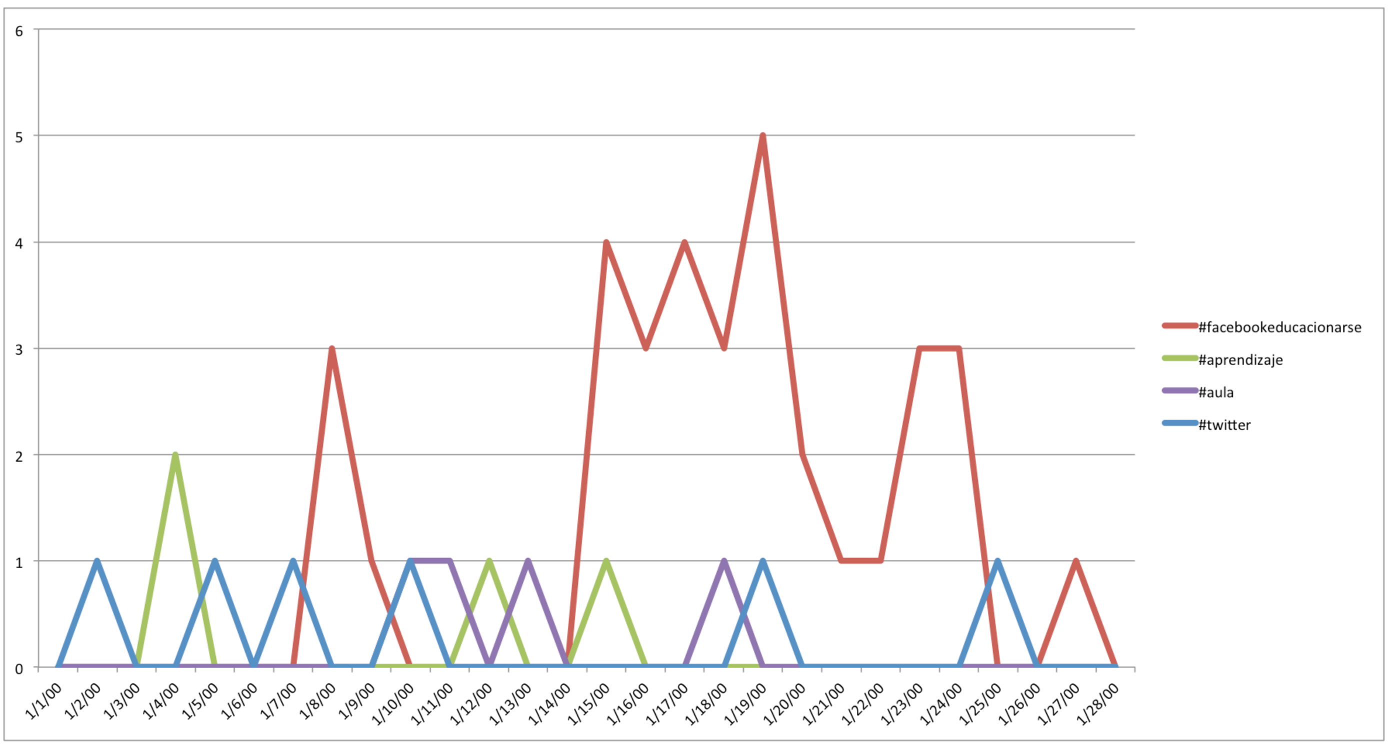


Figure 13. Evolution of the use of informal hashtags on Twitter throughout the course.

IV. Discussion and Conclusions

Both hashtags on Twitter or Google+ and the categories in the latter network represent an opportunity for collaborative MOOCs. Also are an opportunity to those MOOC with characteristics related to non-formal and informal learning, since they allow to classify and collect the contents generated in the communities related to the course, and to feed back the MOOC on the basis of this non-formal and informal content. The main drawback in retrieving this labeling knowledge, as can be observed in the results, is the lack of digital skills, custom, and awareness on the part of the participants in this type of learning communities. Depending on what is observed, usually the result of the conversations are orphaned publications in terms of labels and even in many cases finding that these labels do not match the ones in the course or are poorly written, making it difficult to recover and forcing cleaning tasks and manual selection of publications.

Regarding the survey conducted to the participants of the MOOC, it should be noted that the population over 50 years old is that uses less, and even acknowledged never do, compared to people aged 26-50. Regarding this it is remarkable that those who are known as digital natives [65] and who are more familiar with these technologies are not the ones who use them most. On the other hand there is not a high dispersion in the results regarding the perception of utility of the hashtags, being this very positive and although with little difference, the digital natives group is the one that less utility sees in the use of hashtags. Finally, also related to perception, those users who have not finished the course are those who, by far, find the use of hashatgs less useful.

On the other hand, it is clear how the Google+ network has made it possible to create a differentiated space for the course community through the Communities tool. Thanks to this separation is possible to easily recover conversations even when users utilize their own hashtags and not those defined in the course or directly without using them. On the contrary, in Twitter it seems essential to use in the tweets associated with the course some hashtag (non formal) previously defined to detect the conversations in this way, otherwise, information is lost and it becomes very difficult (although not impossible , according to Figures 12 and 13) recovery of “informal” hashtags. So it is possible to say that Twitter makes difficult the recovery of informal conversations unlike Google+.

Following on from the results, can be remarkable an informal conversation within the community initiated by the students themselves and especially associated to the contribution of content in the form of publication; consequence of the non-formal activity initiated or even infused by the mechanics of the course. It is possible to verify how, for example, as a result of non-formal activities that are expressed as student publications, a totally voluntary interaction of the rest of peers occurs through approval indicators such as “+ 1s”, appearing more than 300 throughout the course associated with such non-formal activities or even associated comments.

On the results that compare the MOOC's ratings with performance in social networks, it is possible to affirm that in many cases, users who indicate their social network and who post messages in them have a greater interest in completing the course. There are considerable results that indicate a relationship between activities in both directions, although it is true that this relationship is more pronounced in the social network Twitter than in Google+.

As a result of these results, and in conclusion, it is possible to affirm that it is possible to recover and classify non-formal and informal learning that students perform in environments such as social networks and that this knowledge can shed light on the complex learning processes that are given in multitudinous digital societies such as the one shown.

As for other possibilities offered by this type of analysis, performing a deeper analysis and user level, it is presumed that it is possible to be able to classify types of users according to their activity in social networks and MOOC (distribution over time, interest indicators, possibility of increasing the segmentation for the adaptability of the MOOC platform through the data observed in Figures 12 and 13, etc.). This would allow researchers to find influential users, users who behave as spectators, users who do not really have an interest in completing a course, but learning from the process, etc.

ACKNOWLEDGMENT

The authors thank the GRIAL and LITI Research Groups for their collaboration in this research work. Similarly, the authors would like to thank the Universidad Politécnica de Madrid for their support and funding through the project “Diseño y desarrollo de MOOC universitarios” (Ref. PT1415-05000) and to the Vice-President for Academic Policy at the University of Salamanca for financing the Educational Innovation project ID2014/0281 that has allowed the authors to develop the software architecture described in this work. Also, the author Juan Cruz-Benito thanks the European Social Fund and the Junta de Castilla y León (Spain) for funding his pre-doctoral fellow contract.

REFERENCES

[1] F. J. García-Peñalvo and A. M. Seoane Pardo, "Una revisión actualizada del concepto de eLearning. Décimo Aniversario," *Education in the Knowledge Society (EKS),* vol. 16, pp. 119-144, 2015.

[2] B. A. Collis, *Tele-learning in a digital world: The future of distance learning*: International Thomson Computer Press, 1996.

[3] F. J. García-Peñalvo, *Advances in E-Learning: Experiences and Methodologies*. Hershey, PA, USA: Information Science Reference (formerly Idea Group Reference), 2008.

[4] F. J. García-Peñalvo, "Docencia," in *Libro Blanco de la Universidad Digital 2010* J. Laviña Orueta and L. Mengual Pavón, Eds., ed Barcelona, España: Ariel, 2008, pp. 29-61.

[5] M. J. Rosenberg, *E-learning: Strategies for Delivering Knowledge in the Digital*. New York, NY, USA: McGraw-Hill, 2001.

[6] F. Llorens, "Campus virtuales: de gestores de contenidos a gestores de metodologías," *RED. Revista de Educación a Distancia,* pp. 1-12, 2014.

[7] J.-M. Dodero, F.-J. García-Peñalvo, C. González, P. Moreno-Ger, M.-A. Redondo, A. Sarasa*, et al.*, "Development of E-Learning Solutions: Different Approaches, a Common Mission," *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje (IEEE RITA),* vol. 9, pp. 72-80, 2014.

[8] F. Llorens, R. Molina, P. Compañ, and R. Satorre, "Technological Ecosystem for Open Education," in *Smart Digital Futures 2014. Frontiers in Artificial Intelligence and Applications*. vol. 262, R. Neves-Silva, G. A. Tsihrintzis, V. Uskov, R. J. Howlett, and L. C. Jain, Eds., ed: IOS Press, 2014, pp. 706-715.

[9] A. J. Berlanga and F. J. García-Peñalvo, "Learning Design in Adaptive Educational Hypermedia Systems," *Journal of Universal Computer Science,* vol. 14, pp. 3627-3647, 2008.

[10] P. Brusilovsky, "Adaptive Hypermedia," *User Modeling and User-Adapted Interaction,* vol. 11, pp. 87-110, 2001.

[11] A. J. Berlanga and F. J. García-Peñalvo, "Learning Technology Specifications: Semantic Objects for Adaptive Learning Environments," *International Journal of Learning Technology,* vol. 1, pp. 458-472, 2005.

[12] K. M. Kapp, *The gamification of learning and instruction: game-based methods and strategies for training and education*. San Francisco, CA, USA: Wiley, 2012.

[13] C. González-González and F. Blanco-Izquierdo, "Designing social videogames for educational uses," *Computers & Education,* vol. 58, pp. 250-262, 2012.

[14] G. Surendeleg, U. Tudevdagva, and Y. S. Kim, "The Contribution of Gamification on User Engagement in Fully Online Course," in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RITA.2017.2655218, IEEE Revista Iberoamericana de Technologias del Aprendizaje

Creativity in Intelligent, Technologies and Data Science: First Conference, CIT&DS 2015, Volgograd, Russia, September 15–17, 2015, Proceedings, A. Kravets, M. Shcherbakov, M. Kultsova, and O. Shabalina, Eds., ed Cham: Springer International Publishing, 2015, pp. 710-719.

[15] Á. Fidalgo-Blanco, M. L. Sein-Echaluce, F. J. García-Peñalvo, and M. Á. Conde, "Using Learning Analytics to improve teamwork assessment," *Computers in Human Behavior,* vol. 47, pp. 149-156, 2015.

[16] J. Cruz-Benito, R. Therón, F. J. García-Peñalvo, and E. Pizarro Lucas, "Discovering usage behaviors and engagement in an Educational Virtual World," *Computers in Human Behavior,* vol. 47, pp. 18-25, 6// 2015.

[17] M. Á. Conde and Á. Hernández-García, "Learning analytics for educational decision making," *Computers in Human Behavior,* vol. 47, pp. 1-3, 6// 2015.

[18] F. G. Martin, "Will massive open online courses change how we teach?," *Communications of the ACM,* vol. 55, pp. 26-28, 2012.

[19] F. J. García-Peñalvo, Á. Fidalgo Blanco, and M. L. Sein-Echaluce Lacleta. (2014, 6/1/2014). *Tendencias en los MOOCs*. Available: http://gredos.usal.es/jspui/handle/10366/125093

[20] J. I. Aguaded Gómez, "La revolución MOOCs, ¿una nueva educación desde el paradigma tecnológico?," *Comunicar,* pp. 7-8, 2013.

[21] C. Castaño Garrido, I. Maiz Olazabalaga, and U. Garay Ruiz, "Diseño, motivación y rendimiento en un curso MOOC cooperativo," *Comunicar,* vol. 44, pp. 19-26, 2015.

[22] A. Chiappe-Laverde, N. A. Hine, and J. A. M. Silva, "Literatura y práctica: una revisión crítica acerca de los MOOC," *Comunicar,* pp. 9-18, 2015.

[23] V. J. Marsick and K. E. Watkins, "Informal and Incidental Learning," *New Directions for Adult and Continuing Education,* vol. 2001, pp. 25-34, 2001.

[24] F. J. García-Peñalvo, R. Colomo-Palacios, and M. D. Lytras, "Informal learning in work environments: training with the Social Web in the workplace," *Behaviour & Information Technology,* vol. 31, pp. 753-755, 2012.

[25] H. Eshach, "Bridging in-school and out-of-school learning: Formal, non-formal, and informal education," *Journal of science education and technology,* vol. 16, pp. 171-190, 2007.

[26] Á. Fidalgo Blanco, M. L. Sein-Echaluce Lacleta, and F. J. García-Peñalvo, "Methodological Approach and technological Framework to break the current limitations of MOOC model," *Journal of Universal Computer Science,* vol. 21, pp. 712-734, 2015.

[27] L. Lane. (2012, 10/1/2015). *Three Kinds of MOOCs.* . Available: http://lisahistory.net/wordpress/2012/08/three-kinds-of-moocs/

[28] D. Clark. (2013, 10/1/2015). *MOOCs: taxonomy of 8 types of MOOC.* . Available: http://donaldclarkplanb.blogspot.com.es/2013/04/moocs-taxonomy-of-8-types-of-mooc.html

[29] G. Conole, "Los MOOCs como tecnologías disruptivas: estrategias para mejorar la experiencia de aprendizaje y la calidad de los MOOCs," *Campus Virtuales. Revista Científica Iberoamericana de Tecnología Educativa,* vol. 2, pp. 16-28, 2013.

[30] F. J. García-Peñalvo, D. Griffiths, M. Johnson, P. Sharples, and D. Sherlock, "Problems and opportunities in the use of technology to manage informal learning," in *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality*, F. J. García-Peñalvo, Ed., ed New York, USA: ACM, 2014, pp. 573-580.

[31] N. A. Christakis and J. H. Fowler, *Connected: The surprising power of our social networks and how they shape our lives*. New York, USA: Little, Brown and Company, 2009.

[32] C. Alario-Hoyos, M. Pérez-Sanagustín, C. Delgado-Kloos, H. Parada G, M. Muñoz-Organero, and A. Rodríguez-de-las-Heras, "Analysing the Impact of Built-In and External Social Tools in a MOOC on Educational Technologies," in *Scaling up Learning for Sustained Impact*. vol. 8095, D. Hernández-Leo, T. Ley, R. Klamma, and A. Harrer, Eds., ed: Springer Berlin Heidelberg, 2013, pp. 5-18.

[33] J. Sanchez, C. González, S. Alayón, and P. Gonzalez, "Using social networks at university: The case of school of computer science," in *Global Engineering Education Conference (EDUCON), 2013 IEEE*, ed USA: IEEE, 2013, pp. 492-496.

[34] J. Mackness, S. Mak, and R. Williams, "The ideals and reality of participating in a MOOC," in *7th International Conference on Networked Learning*, ed, 2010, pp. 266-275.

[35] A. McAuley, B. Stewart, G. Siemens, and D. Cormier. (2010). *The MOOC model for digital practice*. Available: http://www.elearnspace.org/Articles/MOOC_Final.pdf

[36] F. J. García-Peñalvo, M. Johnson, G. R. Alves, M. Minović, and M. Á. Conde-González, "Informal learning recognition through a cloud ecosystem," *Future Generation Computer Systems,* vol. 32, pp. 282-294, 2014.

[37] K. Silius, T. Miilumaki, J. Huhtamaki, T. Tebest, J. Merilainen, and S. Pohjolainen, "Students' motivations for social media enhanced studying and learning," *Knowledge Management & E-Learning: An International Journal (KM&EL),* vol. 2, pp. 51-67, 2010.

[38] G. Siemens, "Connectivism: A learning theory for the digital age," *International journal of instructional technology and distance learning,* vol. 2, pp. 3-10, 2005.

[39] M. Zapata-Ros, "Teorías y modelos sobre el aprendizaje en entornos conectados y ubicuos. Bases para un nuevo modelo teórico a partir de una visión crítica del "conectivismo"," *Education in the Knowledge Society (EKS),* vol. 16, pp. 69-102, 2015.

[40] C. Evans, "Twitter for teaching: Can social media be used to enhance the process of learning?," *British Journal of Educational Technology,* vol. 45, pp. 902-915, 2014.

[41] E. M. M. Gallardo Echenique, Luis and M. Bullen, "Students in higher education: Social and academic uses of digital technology," *RUSC. Universities and Knowledge Society Journal,* vol. 12, pp. 25-37, 2015.

[42] E. C. Wenger and W. M. Snyder, "Communities of practice: The organizational frontier," *Harvard business review,* vol. 78, pp. 139-146, 2000.

[43] J. West, "Recognition of non-formal and informal learning: the Case Against. Study prepared for the meeting of the OECD Group of Experts," Vienna. 2007.

[44] J. Cruz-Benito, O. Borrás-Gené, F. J. García-Peñalvo, Á. Fidalgo Blanco, and R. Therón, "Detección de aprendizaje no formal e informal en Comunidades de Aprendizaje soportadas por Redes Sociales en el contexto de un MOOC Cooperativo," in *Actas del XVII Simposio Internacional de Informática Educativa (SIIE'15) (Setúbal, Portugal. 25-27 November 2015)*, M. d. R. r. Rodrigues, M. n. Llamas Nistal, and M. Figueiredo, Eds., ed Portugal: Escola Superior de Educação do Instituto Politécnico de Setúbal, 2015, pp. 410-418.

[45] J. Cruz-Benito, O. Borrás-Gené, F. J. García-Peñalvo, Á. Fidalgo Blanco, and R. Therón, "Detection of Non-Formal and Informal Learning in learning communities supported by social networks in the context of a Cooperative MOOC," in *Proceedings of the XVII International Symposium on Computers in Education (SIIE'15) (Setúbal, Portugal. 25-27 November 2015)*, M. d. R. r. Rodrigues, M. n. Llamas Nistal, and M. Figueiredo, Eds., ed Portugal: IEEE, 2015.

[46] O. Borrás-Gené, M. Martínez-Núñez, and Á. Fidalgo-Blanco, "New Challenges for the Motivation and Learning in Engineering Education Using Gamification in MOOC," *International Journal of Engineering Education (IJEE),* vol. 32, pp. 501-512, 2016.

[47] D. Clow, "MOOCs and the funnel of participation," presented at the Proceedings of the Third International Conference on Learning Analytics and Knowledge, Leuven, Belgium, 2013.

[48] M. Waite, J. Mackness, G. Roberts, and E. Lovegrove, "Liminal participants and skilled orienteers: Learner participation in a MOOC for new lecturers," *Journal of Online Learning and Teaching,* vol. 9, p. 200, 2013.

[49] F. J. García-Peñalvo, J. Cruz-Benito, O. Borrás-Gené, and Á. Fidalgo Blanco, "Evolution of the Conversation and Knowledge Acquisition in Social Networks Related to a MOOC Course," in *Learning and Collaboration Technologies*. vol. 9192, P. Zaphiris and A. Ioannou, Eds., ed: Springer International Publishing, 2015, pp. 470-481.

[50] Universidad Politécnica de Madrid, Universidad de Zaragoza, and Universidad de Salamanca. (2015). *iMOOC*. Available: http://gridlab.upm.es/imooc/

[51] M. L. Sein-Echaluce Lacleta, Á. Fidalgo Blanco, F. J. García-Peñalvo, and M. Á. Conde-González, "iMOOC Platform: Adaptive

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RITA.2017.2655218, IEEE Revista Iberoamericana de Technologias del Aprendizaje

MOOCs," in *Learning and Collaboration Technologies*, ed: Springer International Publishing, 2016, p. In press.

[52] Á. Fidalgo Blanco, F. J. García-Peñalvo, and M. L. Sein-Echaluce Lacleta, "A methodology proposal for developing adaptive cMOOC," in *Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality*, F. J. García-Peñalvo, Ed., ed New York, USA: ACM, 2013, pp. 553-558.

[53] N. Sonwalkar, "The first adaptive MOOC: A case study on pedagogy framework and scalable cloud Architecture—Part I," *MOOCs Forum,* vol. 1, pp. 22-29, 2013.

[54] J. Daniel, E. V. Cano, and M. Gisbert, "The Future of MOOCs: Adaptive Learning or Business Model?," *RUSC. Universities and Knowledge Society Journal,* vol. 12, pp. 64-73, 2015.

[55] Á. Fidalgo Blanco, M. L. Sein-Echaluce Lacleta, and F. J. García-Peñalvo, "MOOC cooperativo. Una integración entre cMOOC y xMOOC. Cooperative MOOC. An integration between cMOOC and xMOOC," in *II Congreso Internacional sobre Aprendizaje, Innovación y Competitividad, CINAIC 2013*, Á. F. Blanco and M. L. S.-E. Lacleta, Eds., ed Madrid: Fundación General de la Universidad Politécnica de Madrid, 2013, pp. 481-486.

[56] Á. Fidalgo Blanco, M. L. Sein-Echaluce Lacleta, O. Borrás Gené, and F. J. García Peñalvo, "Educación en abierto: Integración de un MOOC con una asignatura académica," *Education in the Knowledge Society (EKS),* vol. 15, pp. 233-255, 2014.

[57] J. Mackness, M. Waite, G. Roberts, and E. Lovegrove, "Learning in a small, task–oriented, connectivist MOOC: Pedagogical issues and implications for higher education," *2013,* vol. 14, 2013-09-30 2013.

[58] O. Borrás Gené, M. Martínez Núñez, and Á. Fidalgo Blanco, "Gamification in MOOC: challenges, opportunities and proposals for advancing MOOC model," in *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality*, F. J. García-Peñalvo, Ed., ed New York, USA: ACM, 2014, pp. 215-220.

[59] Á. Fidalgo-Blanco, M. L. Sein-Echaluce, F. J. García-Peñalvo, and J. Esteban Escaño, "Improving the MOOC learning outcomes throughout informal learning activities," in *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality*, F. J. García-Peñalvo, Ed., ed New York, USA: ACM, 2014, pp. 611-617.

[60] J. Huang, K. M. Thornton, and E. N. Efthimiadis, "Conversational tagging in twitter," in *Proceedings of the 21st ACM conference on Hypertext and hypermedia*, 2010, pp. 173-178.

[61] M. Efron, "Hashtag retrieval in a microblogging environment," in *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, 2010, pp. 787-788.

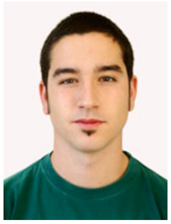
[62] J. Cruz-Benito, O. Borrás-Gené, F. J. García-Peñalvo, Á. Fidalgo Blanco, and R. Therón, "Extending MOOC ecosystems using web services and software architectures," in *Proceedings of the XVI International Conference on Human Computer Interaction*, ed Vilanova i la Geltrú, Spain: ACM, 2015, pp. 1-7.

[63] J. Cruz-Benito, R. Therón, and F. J. García-Peñalvo, "Software Architectures Supporting Human-Computer Interaction Analysis: A Literature Review," in *Learning and Collaboration Technologies: Third International Conference, LCT 2016, Held as Part of HCI International 2016, Toronto, ON, Canada, July 17-22, 2016, Proceedings*, P. Zaphiris and A. Ioannou, Eds., ed Cham: Springer International Publishing, 2016, pp. 125-136.

[64] J. Cruz-Benito, O. Borrás-Gené, F. J. García-Peñalvo, Á. Fidalgo Blanco, and R. Therón, "Comunidades de Aprendizaje en Redes Sociales y su Relación con los MOOC," *VAEP RITA,* vol. 4, pp. 87-99, 2016.

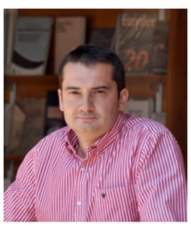
[65] F. García, J. Portillo, J. Romo, and M. Benito, "Nativos digitales y modelos de aprendizaje," in *SPDECE*, 2007.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RITA.2017.2655218, IEEE Revista Iberoamericana de Technologias del Aprendizaje



Juan Cruz-Benito received the M.Sc. degree in intelligent systems from the University of Salamanca, Spain, in 2013, where he is currently pursuing the Ph.D. degree in computer sciences. He is one of the youngest members of the Research Group Interaction and eLearning, where he specializes in software solutions based on technology ecosystems. He is an expert in educational vir-tual worlds, data analytics, and technologies for educational purposes, and disciplines that he has developed in many innovation and research projects. He has participated in many European and national projects, such as Tagging, Recognition and Acknowledgment of Informal Learning Experiences, Virtual Alliances for Learning Society, USALSIM Virtual Campus, or the Spanish Observatory for University Employability and Employment, where he participated as a Software Engineer, Researcher, and Developer.

Oriol Borrás-Gené Is a Telecommunications Engineer and PhD student in Systems Engineering and Services for the Information Society by the Universidad Politécnica de Madrid, an expert in social networks, gamification and MOOC. He works as an eLearning technician in the Tele-Education Office of the UPM as part of different projects such as collaborative networking, virtual laboratories and MOOC. He is a teacher-coordinator of the course "Application of social networks to teaching: Virtual Communities"



Francisco José García-Peñalvo received the degrees in computing from the University of Salamanca and the University of Valladolid, and the Ph.D. degree from the University of Salamanca. He is currently the Head of the Research Group Interaction and eLearning. He has led and partici-pated in over 50 research and innovation projects. He was the Vice Chancellor of Innovation with the University of Salamanca from 2007 to 2009. He has authored over 300 articles in international journals and conferences. His main research interests focus on eLearning, computers and education, adaptive systems, Web engineering, semantic Web, and software reuse. He was a Guest Editor of several special issues of international journals, such as *Online Information Review*, *Computers in Human Behavior*, and *Interactive Learning Environments*. He is also a member of the Program Committee of several international conferences and a Reviewer of several international journals. He is currently the Editor-in- Chief of the *International Journal of Information Technology Research* and the *Education in the Knowledge Society Journal*. Besides, he is the Coordinator of the multidisciplinary Ph.D. Programme on Education in the Knowledge Society.

Ángel Fidalgo Blanco holds a PhD in Computer Science from the University of Las Palmas de Gran Canaria and is a professor at the ETSI of Mines and Energy at the Universidad Politécnica de Madrid. He leads the Laboratory of Innovation in Technologies of Innovation of the Department of Geological and Mining Engineering of the Universidad Politécnica de Madrid (UPM). He is also president of the organizing committee of the International Congress on Learning, Innovation and Competitiveness (CINAIC). He received the "Award for the best OCW subject" awarded by the Ministry of Education of the Government of Spain and the UPM award for Educational Innovation. He is the editor of the blog "Innovación Educativa" (https://innovacioneducativa.wordpress.com/) and creator of the learning community "connectivismo.net" (https://goo.gl/vk6wzr). He has been working in R+D+i on educational innovation for more than 25 years, generating innovative products and applications based on adaptive learning, knowledge management, cooperative work and new training models. Committed to dissemination, it regularly participates in courses, seminars and conferences on educational innovation.

Roberto Therón studied Informatics at the University of Salamanca (Diploma) and the University of La Coruña (BA). After joining the Research Group Robotics at the University of Salamanca, he presented his thesis work, "Parallel calculation configuration space for redundant robots", receiving the Extraordinary Doctoral Award. He subsequently obtained the degrees of Bachelor in Communication Studies (University of Salamanca) and Bachelor in Humanities (University of Salamanca). At the same University of Salamanca he continues to carry out his research work, as manager of VisUsal group (within the Recognized Research Group GRIAL) which focuses on the combination of approaches from computer science, statistics, Graphic Design and Information Visualization for obtain an adequate understanding of complex data sets. In recent years, he has been dedicated to developing advanced visualization tools for multidimensional data, such as genetic or paleo-climate data. In the field of Visual Analytics he develops productive collaborations with groups and institutions internationally recognized as the Laboratory of Climate Sciences and the Environment (Paris) or the Center for Advanced Visual Analytics UN (Switzerland). He is the author of over 70 articles in international journals and conferences.

6.8 Appendix H. Improving success/completion ratio in large surveys: a proposal based on usability and engagement

Cruz-Benito, J., Therón, R., García-Peñalvo, F. J., Sánchez-Prieto, J. C., Vázquez-Ingelmo, A., Martín-González, M., & Martínez, J. M. (2017). Improving success/completion ratio in large surveys: a proposal based on usability and engagement. In P. Zaphiris & A. Ioannou (Eds.), *Learning and Collaboration Technologies. Technology in Education. 4th International onference, LCT 2017. Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9–14, 2017. Proceedings, Part II* (pp. 352-370). Switzerland: Springer International Publishing. doi:10.1007/978-3-319-58515-4_28

The final publication is available at link.springer.com

Improving success/completion ratio in large surveys: a proposal based on usability and engagement

Juan Cruz-Benito[1], Roberto Therón[1], Francisco J. García-Peñalvo[1], José Carlos Sánchez-Prieto[1], Andrea Vázquez-Ingelmo[1], Martín Martín-González[2], Jorge M. Martínez[2]

[1]GRIAL Research Group. Department of Computers and Automatics. University of Salamanca
[2]UNESCO Chair in University Management and Policy. Technical University of Madrid

{juancb, theron, fgarcia, josecarlos.sp, andreavazquez}@usal.es, {martin.martin, jorge.martinez}@upm.es

Abstract. This paper presents a research focused on improving the success/completion ratio in large surveys. In our case, the large survey is a questionnaire produced by the Spanish Observatory for University Employability and Employment (OEEU in the Spanish acronym). This questionnaire is composed by around 32 and 60 questions and between 86 and 181 variables to be measured. The research is based on the previous experience of a past questionnaire proposed by the OEEU composed also by a large amount of questions and variables to be measured (63-92 questions and 176-279 variables). After analyzing the target population of the questionnaire (with the target population of the previous questionnaire as reference) and reviewing the literature, we have designed 11 proposals for changes in the questionnaire that could improve users' completion and success ratios (changes that could improve the users' trust in the questionnaire, the questionnaire usability and user experience or the users' engagement to the questionnaire). These changes are planned to be applied in the questionnaire in two main different experiments based on A/B test methodologies that will allow researchers to measure the effect of the changes in different populations and in an incremental way. The proposed changes have been assessed by five experts through an evaluation questionnaire. In this questionnaire, researchers gathered the score of each expert regarding to the pertinence, relevance and clarity of each change proposed. Regarding the results of this evaluation questionnaire, the reviewers fully supported 8 out of the 11 changes proposals, so they could be introduced in the questionnaire with no variation. On the other hand, 3 of the proposed changes or improvements are not fully supported by the experts (they have not received a score in the top first quartile of the 1-7 Likert scale). These changes will not be discarded immediately, because despite they have not received a Q1 score, they received a score within the second quartile, so could be reviewed to be enhanced to fit the OEEU's context.

Keywords. Human-Computer Interaction, HCI, Online Survey, Online Questionnaire, Usability, User Experience, Engagement, Trust, A/B Test.

1 Introduction

The collection of information by questionnaires and interviews is one of the most well-known and currently used methods to get users' opinions, both in the physical and digital environments.

It is common in many websites to have a form for entering information, either as a contact point, as part of the login for the system, as part of a payment process, etc. The forms are so integrated into the web user interaction, that their importance is relativized and it is assumed that the user will complete it by the mere fact that they are faced to them regularly. However, this is not so.

Indeed, the web forms pervasivity, in recent years have triggered certain trends and user behaviors towards such information entry tools. For example, it has been proven [1] the following regarding users' behavior towards forms:

- Users rely more on websites, even being more willing to perform complex actions (at all levels), such as purchases, payments, etc.
- They protect more their information, they are less willing to disclose personal information.
- They demand better products, are less tolerant to bad forms.

During the last years a lot of work has been carried out in relation to the questionnaires, establishing that users have some reluctance to complete a form from even before to begin filling it [1]. This poses certain problems regarding the achievement of information collection objectives intrinsic to any form.

Regarding the types of users who complete forms, different profiles can be set [1]:

1. Readers: Those who read the form carefully.
2. Rushers: These users rush in and begin completing fields, reading only when they think it is necessary.
3. Refusers: These users won't have anything to do with the form.

According to the literature, and intimately related to the Social Exchange Theory [2], some authors [1] distinguish three layers in the forms: relationship, conversation and appearance.

1. The relationship of a form is based on the relationship that who asks the questions has with whom responds.
2. The conversation of a form goes from the questions that are asked, to the instructions given or to the organization of the questions according to their topic.

3. The appearance of the form is the image it displays: placement of text, graphics, areas of data entry, color, etc.

Improving these factors, such as the relationship with the user, makes it easier for the user to participate and complete his task within the questionnaire.

This paper presents a research aimed at designing and validating different changes in the context of a very large questionnaire regarding users' trust, user experience, usability and engagement with the final goal of improving the users' completion/success ratios. These possible improvements are compared with another questionnaire previously developed for the same topics and context, by means of different methodologies and approaches. To present this research, the paper will have the following structure: section 2 provides the needed context of the questionnaires and population study; section 3 presents the research goals and experiments design; the fourth section comments on the proposed changes and improvements designed by the researchers; the fifth section presents the evaluation of the proposals carried out by experts. Finally, the sixth section presents the conclusions of the paper and outlines the future work to be done regarding this research.

2 Background: The Spanish Observatory for University Employability and Employment (OEEU)

During the months of June-July 2015, the Spanish Observatory for University Employability and Employment (OEEU) contacted several thousand Spanish university graduates (133588 individuals) through the universities (48, public and private institutions) where they got their degrees in the course 2009-2010 to invite them to fill out a questionnaire [3, 4].

This questionnaire had a common part with 60 questions and 167 variables measured, in addition to 3 specific itineraries depending on the users' previous responses. The first itinerary added 3 questions and 9 measured variables more. The second one, added 24 questions and 70 variables. Finally, the third itinerary added 32 more questions and 112 variables to the common part of the questionnaire.

Therefore, the questionnaire varies between 63-92 questions and 176-279 variables depending on the itinerary that the user follows. It can be stated without doubt that the questionnaire is very extensive.

The number of users who started the questionnaire was 13006 (9.74% of the total population), of which 9617 completed it (7.20% of the total population, 73.94% of the total started questionnaires).

The descriptive data regarding the age of the participants in the questionnaire were the following (the count of users is 12109 because the birthdate data was not mandatory and not all users filled it out):

```
count   12109.000000
mean       32.525972
std         7.018282
min        25.000000
```

25%	28.000000
50%	30.000000
75%	34.000000
max	80.000000

As for gender, 56.05% (7290) of the users who answered the questionnaire were women and 43.94% (5716) men. In relation to nationality, 98.54% (11672) of the users were Spanish and 1.46% (173) were foreigners.

About the users who dropped out of the questionnaire, the quartiles of the dropout rate based on the questionnaire screen where they left off were:

count	3389.000000
25%	4.000000
50%	5.000000
75%	7.000000

That is, 25% of the users left on screen 4 or before, another 25% left between screens 4 and 5 of the questionnaire, another 25% between screens 5 and 7 and another 25% between screen 7 and the end (depending on the itinerary).

Now, in 2017, a process of gathering information similar to the one carried out during 2015 will conducted again. In this case, the information to be collected is about graduates of masters studies that ended their studies during the 2013-2014 academic year. For this purpose, a questionnaire composed of between 32 and 60 questions and between 86 and 181 variables to be measured has been proposed (the questionnaire has again several itineraries depending on the user's answers). Without too much analysis in detail, it can be considered that despite the differences, it is a large questionnaire and shares some of the problems of the previous one in terms of difficulties or challenges that can appear during its completion by the users.

Before sending out the questionnaires to the students, the Observatory gathers some data about students from the participant universities. Currently, on February 2017, there are collected data from 28744 people coming from 32 public and private Spanish universities. About these former students, the Observatory have the following data:

Descriptive data regarding the age of the population to which the questionnaire will be addressed:

count	28744.000000
mean	35.854370
std	15.852381
min	5.000000
25%	28.000000
50%	31.000000
75%	38.000000
max	117.000000

Regarding the data about the age, obviously, the aging of the population with respect to the one of the previous questionnaire is noticed. This is normal taking into account that the required age to begin a master degree is higher than that required to access to a degree (at least on a regular basis).

Regarding the gender of the population to which the questionnaire will be addressed, 55.2% (16385) are women and 44.8% (13317) are men. In relation to nationality, this is the aspect in which the current population (graduates from master degree) is more differentiated from the study performed with degree graduates. This time, the proportion of foreign students is greater, with 88.11% (25318) of Spanish students compared to 11.88% (3414) students with foreign nationality.

In general terms, it is possible to assume that populations (putting each of them in context) are not very different. In this sense, can be highlighted the main difference is in terms of nationality. This difference could lead to consider treating differently aspects of the questionnaire to adjust to possible cultural differences. In this case, there will be no cultural distinction when designing, presenting or performing the questionnaire. This could be considered a limitation of the study.

3 Research goals and experiments proposal

3.1 Overall research goals

The main goals of the experiment that is being designed, and that will be presented below, are:

- Study how to improve the ratio participants actually starting the questionnaire (previously, close to 9.7% of the total population).
- Study how to improve the completion rate of the questionnaire (previously 73.94% completion rate).

In addition to these fundamental goals, another objective related to the second one can be proposed; namely, to grant that in case of dropout, users have completed all possible screens of the form (obtaining by this way more information even if they leave it).

3.2 Experiments proposal

For the new version of the questionnaire, it is considered that several points can be improved compared to the questionnaire implemented in 2015 and to the ways of increasing users' participation.

To implement these improvements, it is proposed to carry out two experiments in parallel:

- A study on how to improve the invitation to graduates processes.
- A study on what improvements can be implemented in questionnaires to improve participation and completion ratios.

The key aspects of each of the studies will be discussed below, indicating the main changes to be implemented, etc.

Also, before implementing these changes to the questionnaires, in addition to being supported in part by the literature, they have been subject of experts validation through a questionnaire [5].

3.2.1 *Study about how to improve the invitation processes to graduates*

In the case of the questionnaires produced by the OEEU, it is necessary to consider a fundamental factor: the privacy of the user is a primary concern over above all else (among other reasons, due to sensitive data being handled). This project is respectful and complies to Spanish Personal Data Protection Act (LOPD), having registered the OEEU's database by the Spanish authorities to safeguard the data.

Due to the privacy restrictions imposed within the project, the Observatory does not keep data that would allow to relate a person with its information. That is, there is no information related to names, ID, exact date of birth, etc. The only exception is that the Observatory offers the option to users of including their email to get information about the investigations, or the results of the draw of some devices (Android tablets) held among the graduates who complete the questionnaire.

In view of these restrictions, and because of the e-mail -if it is obtained at all occurs at the end of the whole process- the universities are the responsible for contacting their graduates offering them to participate in the process of the questionnaires. In this contact message, universities tell graduates that there is a draw among those completing this form and provide a personal link to each student to complete the task. This invitation letter designed by the OEEU Observatory could be used or not by universities, being responsible each one of them of its use and modification.

The experiment proposal in this respect is based on sending two different invitation letters. One invitation letter will be an updated version of this text used for the previous questionnaires (updated to reflect the changes related to the new edition). The second invitation letter will change both in the textual content and visual appearance, applying some changes that will be explained in following sections of this paper (basically modifying the tone and textual content of the message, plus providing a different overall design to the message [6]).

The goal of these two different invitation letters is to send one (the old version) as invitation letter for most part of the universities. The second one (the new) will be used by universities that participated in the previous edition of the questionnaires phase to test if the changes lead to variations in the entrance and participation in the questionnaire changes over the previous edition. With this proposal, it is possible to see the effect of the changes in the invitation letter (using A/B test methodologies) considering several things:

- The context of each participating university is different (population, economic factors, etc.). Therefore, specific changes are made for universities participating in both calls for data collection.
- The population of the study has changed from the first edition of the data collection to this second (age, training, etc.). For this reason, the changes between different universities will be also validated within the same edition of the questionnaires for data collection.

In the following sections, the changes to be introduced will be discussed in depth. At any rate, these proposed changes that could be introduced in the questionnaires are limited by the various constraints of the project related to privacy (it is not possible to use external mailing platforms, etc.) and they focus fundamentally on the issues of improving trust and relationship between the user and the entity that proposes the questionnaire (the Observatory).

3.2.2 A study about what improvements can be implemented in questionnaires to improve participation and completion ratios

Regarding the part of the study related to the changes in the questionnaire itself, several modifications are proposed at several levels [7].

The general approach of this study is to perform an A/B test with three variants (A/B/C). The proposal is composed by a main variant (A) that follows the outline of the previous edition questionnaire (available in Spanish in http://gredos.usal.es/jspui/bitstream/10366/127374/5/Anexos_OEEU_2015.pdf), from which we have some idea of efficiency, etc., along with two other variants (B and C) that change certain issues related to the Social Exchange Theory [2].

In general, variant B of the test refers to changes related to the relationship of the participant (who answers) and who proposes the questionnaire (first layer of theory) along with changes related to the appearance (third layer of the theory) [8, 9]. More broadly, this variant B is based on trust between the parties [10, 11], further improvements and changes with respect to user experience [12], usability [9] and interface design of the questionnaire [10, 13].

On the other hand, variant C of the test includes the proposed changes in variant B plus other changes related to the relationship between the stakeholders involved in the questionnaire (first layer of the theory) and to the conversation between them (second layer). From this point of view, variant C will focus more on issues related to user engagement [6].

In any case, the three versions of the questionnaire will maintain certain rewards offered in the previous process of data collection. For example, this time there will be again a draw of electronic devices (tablets) among those who complete the questionnaire. Also, the Observatory will continue maintaining communication with those users who want to receive the latest news of the Observatory and its research.

Regarding some factors such as age, disability, or other situations and personal contexts of users, in this case they will be obviated (except the application of general accessibility standards) because the experiment is not focused on specific aspects related to possible subgroups within the population of the study [9, 14]. It is assumed that this constitutes a limitation of the study.

The effect of the changes will be measured in two ways:

- Checking the data regarding the access ratio to the questionnaires, the completion of each part of the questionnaire and the completion ratio of the questionnaire (completed screens, dropout moments, etc.).
- Evaluating the *paradata* [15]. The *paradata* from a questionnaire are the auxiliary data that describe a process, such as response times, clicks, scroll processes, etc. In this case, the *paradata* will be related to the time it takes to the user completing the task of answering each page of questions, the time to complete the full

questionnaire, the accesses to the questionnaire, etc. These *paradata* cannot be compared with similar data from the previous round of data collection about degree graduates, since nothing similar was done in that moment.

Usually, in this kind of research, users complete another questionnaire about their opinion about how they have felt about the questionnaire, how they have been able to solve the task, etc. In this case, due to the length of the questionnaire to be completed and the nature of the project, this research will not be carried out in this way. This is a limitation as to the richness of the results that can be obtained. What researchers plan to do is to invite the students, who decide to give their e-mail voluntarily at the end of the employability and employment questionnaire, to a new specialized questionnaire on these issues.

4 Proposed improvements for the questionnaire

In this section the different improvements designed for the questionnaire are described. The design process has been driven by a literature review. This literature review comprised about 650 books, papers and technical reports. The process for selecting the literature to be reviewed was:

- Making three different queries to the Web of Science and collect the results in order to iterate in reading the titles, abstracts and full content to select those papers really relevant for the topic of this research. The three queries performed were:
  - *(("form*") OR ("questionnaire*") OR ("survey*")) AND "usability" AND "factor*" AND (("web") OR ("online"))*
  - online forms usability
  - (("web" OR "online") AND ("questionnaire?" OR "form?") AND usability)

  This process and its results are gathered in the following spreadsheet https://docs.google.com/spreadsheets/d/1KbOCTVBqKh3Xz5nqqQY9-ywgZ2ggYNldb3OS6SasaXk/edit?usp=sharing. In the spreadsheet the 633 unique results retrieved from the Web of Science and their status regarding to their usage in the research regarding to each review stage are presented.
- Extracting the main references from these papers and books retrieved from the Web of Science and read them. This process lead to review another 15 papers, books, standards and technical reports. Most part of them were used in some way to design the proposals that are explained below.

Once the literature was reviewed, authors designed the improvements and changes for the questionnaire. These improvements and changes are mainly supported or inspired by the literature as well as by ISO usability guidelines and HSS (U.S. Department of Health and Human Services) guidelines [16-20]. The following subsections comment each change and measure, describing for each one its purpose, its goal, the identifier associated, etc.

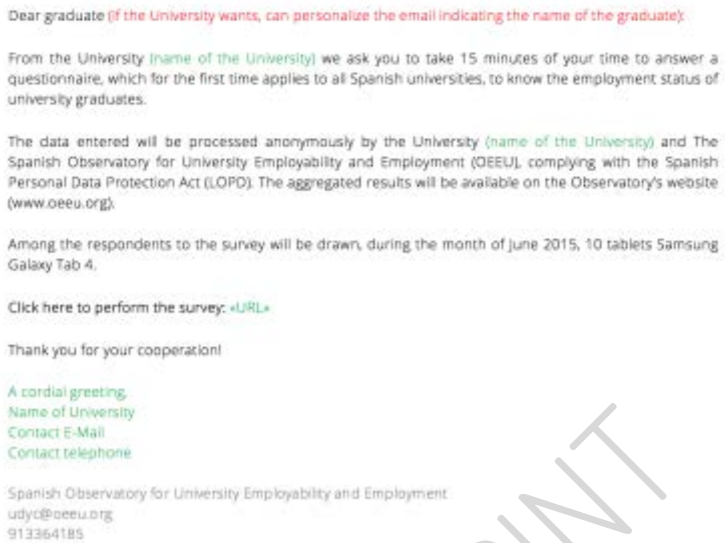
The ID has been set for each proposed change related to the main application area of application within the HCI discipline; despite of that, most of them apply to more than one area, for that reason, researchers pick the main one as base for the identifier. Table 1 explains the relationship between each change/improvement (using the IDs explained in the subsections), its relationship with HCI knowledge areas or topics and with each layer of the Social Exchange Theory used as framework for the experiments design and the research in general. The main improvement areas of each change related to HCI topics are marked in red color and bigger size.

Table 1. Relationship between each change/improvement proposed, HCI application areas and layers of Social Exchange Theory

Layer of the Social Exchange Theory	Relationship	Conversation		Appearance
Improvement area regarding HCI	Trust	Engagement	Usability / User Experience (UX)	Design
TR1	**X**			X
TR2	**X**		X	X
TR3	**X**			
US/UX 1			**X**	X
US/UX 2			**X**	X
US/UX 3	X		**X**	
US/UX 4			**X**	X
TR4	**X**	X		
EN1		**X**		X
EN2		**X**		X
EN3		**X**		

4.1 Proposal for the invitation letter to the questionnaires

Proposed change: *TR1. Modify the text and appearance from the invitation letter to the questionnaire.*



Dear graduate (if the University wants, can personalize the email indicating the name of the graduate):

From the University [name of the University] we ask you to take 15 minutes of your time to answer a questionnaire, which for the first time applies to all Spanish universities, to know the employment status of university graduates.

The data entered will be processed anonymously by the University (name of the University) and The Spanish Observatory for University Employability and Employment (OEEU), complying with the Spanish Personal Data Protection Act (LOPD). The aggregated results will be available on the Observatory's website (www.oeeu.org).

Among the respondents to the survey will be drawn, during the month of June 2015, 10 tablets Samsung Galaxy Tab 4.

Click here to perform the survey: «URL»

Thank you for your cooperation!

A cordial greeting,
Name of University
Contact E-Mail
Contact telephone

Spanish Observatory for University Employability and Employment
udyc@oeeu.org
913364185

Fig. 1. Invitation letter proposed by the OEEU. Text translated by the authors from [5]

In the Figure 1 the basic e-mail, designed by The Spanish Observatory for University Employability and Employment to invite the graduates in the previous edition of the data gathering process, is presented. In this edition of the data gathering process related to master graduates, the basic invitation letter text will be very similar, only changing the text to reflect the master degree of the graduates and specifying that two years ago there were another similar questionnaire that collected data from degree graduates (including also the results displayed in its web http://datos.oeeu.org).

Among the proposed changes are the inclusion of the university logo that sends the invitation, the inclusion of the OEEU logo, a change in design to make the questionnaire according to the colors and fonts used in other OEEU's products, and changes in the text to be perceived as a more personal invitation to the graduate. These changes are intended to improve user trust in the questionnaire and the activity of the Observatory [6, 11, 13, 21].

Figure 2 shows the proposed new design (visual and textual) for the invitation letter. As explained before, the new version will be used only by few universities to allow researchers the measuring of its effect in the graduates.

VNIVERSIDAD
D SALAMANCA

Dear graduate (if the university prefer, can send the mail personalized indicating the name of the graduate):

From the University (name of the University) and the Spanish Observatory for University Employability and Employment (OEEU, http://oeeu.org) we invite you to participate in a study to know the employment situation of the master graduates in Spain. To do this, we ask you to spend a little more than 15 minutes of your time in completing the questionnaire that we propose at the end of this email. In the questionnaire you will be asked about various aspects of your work, personal information (demographic data) and academic record to try to analyze how the master graduates in Spain get employment and what aspects affect their employability. The data you enter will be treated in an absolutely anonymous way and complying with the Spanish Personal Data Protection Act (LOPD).

As a consideration to your collaboration, we will raffle during the month of July 2017, 10 tablets Samsung Galaxy Tab 4. You will also be the first to know the results of those studies conducted with the questionnaire.

Click here to start the survey << Custom URL >>
Thank you very much for your collaboration!

A cordial greeting,
Name of University
Contact E-Mail
Contact telephone

Spanish Observatory for University Employability and Employment
udyc@oeeu.org
913364185

Fig. 2. Invitation letter with visual and textual changes proposed for the research. Adapted from [5]

4.2 Proposal to amend the questionnaire for variant B

Proposed change: *TR2. Adequacy of the image to the other digital products of the Observatory.*

This change is related to modifying colors, logotypes, typography, etc. to correspond the other products of the Observatory like its website http://datos.oeeu.org. This change is supported by the literature as a way to enhance the users' trust in the Observatory brand and products [1, 6, 8, 11, 13, 21].

Proposed change: *TR3. Inclusion of the Observatory's logo and university's logo.*

In the same way that previous proposal, the inclusion of the OEEU logo and the university logotype can reduce the distrust of the graduate to participate. In this case, the logotype of the university will help to build trust on the questionnaire website and the OEEU logotype will help him/her to associate the product with the institution that proposes it [1, 6, 21].

Proposed change: *US/UX1. Inclusion of a progress bar in the questionnaire.*

By observing a progress bar, the user can know its progress in the task of filling the questionnaire and estimate how much effort/time he/she will need to make to complete it. This can reduce the stress related to uncertainty about a task like an unknown questionnaire [1, 6].

Proposed change: *US/UX2. Present a visual focus animation on concrete actions.*

In this case, the web will provide a visual effect of focus to the user in that he will have always in the center of the screen the task to be solved (typically answering a question or filling an empty field), making also a defocused effect on the elements that are not fundamental to solve that task. This proposal is used in commercial questionnaire systems like http://typeform.com/.

The reader can access to the following URLs to check how this visual effect works: https://drive.google.com/file/d/0BwS7cZg3riXtajJtNGhkMnIzXzg/view?usp=sharing, https://drive.google.com/file/d/0BwS7cZg3riXtWGk1bmlvSVB5dDg/view?usp=sharing.

Proposed change: US/UX3. *Deactivation of control elements when an action is initiated.*

A typical example of this change is to deactivate a button in a website once it is pressed until its action is finished. This usability / user experience measure could make the user to trust on the sturdiness of the system and reduce stress situations like those where a button perform the same action several times after being pressed more than once [6].

Proposed change: *US/UX4. In related elements, instead of having smaller and more specific groupings, use some larger grouping, following the Gestalt principles on grouping.*

For example, following the proposal, the header of a table would be fixed while in the content can be scrolled up and down. It seeks to ensure that the large dimensions of analysis in some points of the questionnaire are grouped in an attempt to avoid user fatigue and reducing users' cognitive load when dealing with large tables or complex visual elements [1, 6, 22].

A visual explanation of this proposal can be observed in the following URL https://drive.google.com/file/d/0BwS7cZg3riXtdmZqQzBHZXJVcmM/view?usp=sharing.

4.3 Proposal to amend the questionnaire for variant C

Proposed change: *TR4. Changes in the introduction text to the questionnaire.*

In this case, a change in the text will be sought in a similar way to the modification in the invitation letter to the users. The text changes to a more personal way of addressing the user and contributing important arguments to influence a better perception on what is going to be done and improving the confidence in the questionnaire and the entity that proposes it.

The text of the previous edition is presented in the Figure 3 (the variant A will only update the data about the raffle in the text, etc.).

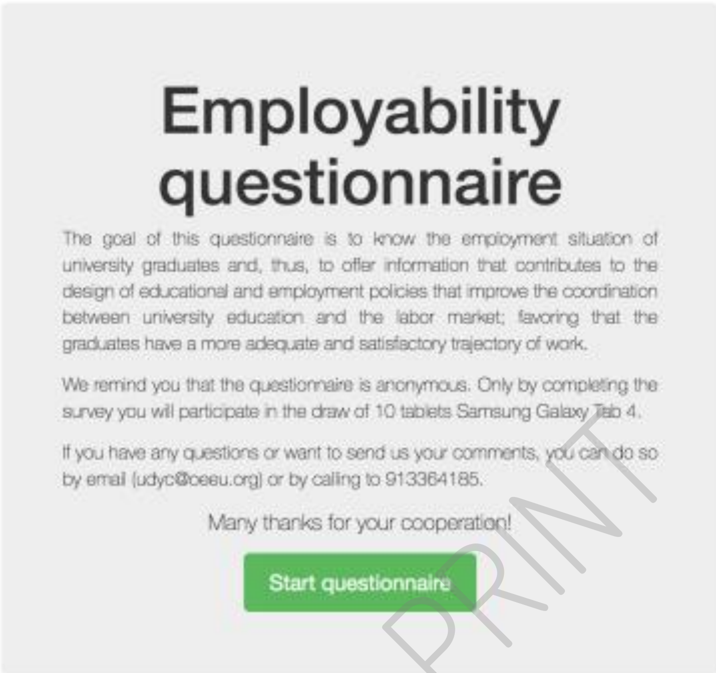


Fig. 3. Previous introduction text to the questionnaire. Translated and adapted from [4, 5]

In this case of variants B and C of the questionnaire, the introductory text would become (changing also the design and layout as commented in the proposal TR2) the displayed in the Figure 4.

Proposed change: *EN1. In the questions related to the community in which they live, change the drop-down selector for a map with the autonomous communities of Spain.*

This will allow the user to select where the user lives through clicking the corresponding one. In this case, it is sought to have visual elements different from the usual ones that allow the user to interact in different ways during the completion of the questionnaire and avoiding to suffer so much fatigue on the repetition of actions. Also, the usage of a map tries to reduce the users' cognitive load that implies the activity of reading a drop-down list of at least 20 items (autonomous communities and cities in Spain). This change is related to some authors that suggest that changing the interaction elements can affect the users' easiness to complete a task [23] and other authors that explain that the time that an user interact with elements in the form is time that the users is not thinking in dropout [6].

Fig. 4. Modification proposal for the introduction text to the questionnaire. Translated and adapted from [5]

Proposed change: *EN2. Inclusion of textual feedback related to user responses including information that may be relevant.*

This inclusion of textual feedback should be placed in at least three different moments of the questionnaire (i.e. after the demographic questions, after the enquiry about whether the graduated has been employed after the master degree or not, and in the final part of the chosen itinerary), regarding the different main dropout moments of the previous data collection process presented in the second section of this paper. This change requires introducing an intermediate screen between two pages of questions in the questionnaire. In this intermediate screen, information in relation to some of his/her answers enabling also comparison of their answers those provided by other users or official stats from other sources, will be provided to the user.

As an example of this kind of feedback, after the screen of the questionnaire where the user responds if he/she has ever worked and how many jobs he/she has had, in the questionnaire screen change (after pressing next), should be displayed a new screen, with only one the the following questions, should be displayed:

- If the graduate answered that he/she did not have a previously a job: "Did you know that there are XX% of graduates in your promotion who have not been able to get a job?".
- If the graduate answered that he/she have had a job: "Do you know that the employment rate of master graduates in Spain is XX%?".
- If the graduate answered that he/she had several jobs: "Did you know what ...? Like you there are XX% of people who have responded to this questionnaire that are in your same situation".

Proposed change: *EN3. Inclusion of web push notifications that allow Observatory to send messages to users in order to encourage them if they leave the questionnaire before finishing.*

These notifications can only be sent if the user explicitly accepts them. The notifications will be accompanied by the link to resume the questionnaire. From a technical point of view, the notifications will be sent to Chrome, Firefox and Safari browsers on Windows, Linux and Mac OS in desktop operative systems and to Android phones with any of those browsers (estimated total market share covered by a 61-77%).

This measure can help to increase the users' engagement as well as to try to improve the completion ratio of the questionnaires through the reinforcement.

Some examples of these kind of notifications are available (in Spanish) in [5].

Also, these web push notifications could help researchers to reach again the participants to invite them to another questionnaire to get feedback about the changes/improvements implemented finally in the form.

5 Evaluation by experts

To validate the proposals designed to improve the questionnaire and reduce the dropout ratio and increase the participation ratio, five experts were invited to evaluate the proposed measures using questionnaire. These experts were selected because all of them work usually with questionnaires from different perspectives (some of them work with questionnaires focusing on improving their usability, use them for research in several contexts, or design questionnaires as part of their day by day work).

In the following subsections, details regarding the questionnaire will be commented, as well, the results and opinions gathered from that questionnaire will be presented and discussed.

5.1 Feedback questionnaire

The assessment questionnaire completed by the experts is based on the proposal by Sánchez-Prieto et al. [24]. In it, the experts assess the relevance of each proposed change, its clarity and its importance, through a Likert scale (1-7 values). In addition the expert can comment on a qualitative way (typing comments in a textbox) any related issues to each question. Also, the questionnaire requires demographic data from the experts related to their gender, knowledge area, etc. [5] in order to characterize them.

5.2 Results and discussion

First, in the validation questionnaire, the experts completed some answers about personal information. In this case, 4 out of 5 experts (80%) were men, 1 (20%) was woman. Regarding the age, 3 out of 5 (60%) are between 41 and 50 years old, while other 2 experts (40%) are between 31 and 40 years old. Regarding their knowledge areas, 3 out of 5 (60%) are related to Engineering and Architecture, while the other 2 (40%) are related to Social and Legal Sciences. Regarding their specialization field, 3 out of 5 (60%) are related to disciplines within Computer Sciences and the other 2 (40%) are related to disciplines within Economics.

Table 2. Descriptive results from the experts' evaluation for each proposal regarding the pertinence, relevance and clarity

	Pertinence			Relevance			Clarity		
	AVG	STD	N	AVG	STD	N	AVG	STD	N
TR1	6.17	0.98	5	5.17	1.17	5	6.17	0.98	5
TR2	6.17	1.33	5	6.00	1.55	5	5.67	1.37	5
TR3	6.67	0.52	5	5.83	1.47	5	6.67	0.52	5
TR4	6.33	1.03	5	6.00	0.89	5	6.50	0.84	5
US/ UX 1	6.50	1.22	5	6.83	0.41	5	6.50	1.22	5
US/ UX 2	7.00	0.00	5	7.00	0.00	5	6.67	0.52	5
US/ UX 3	5.67	2.42	5	6.00	2.45	5	5.33	2.42	5
US/ UX 4	7.00	0.00	5	6.67	0.52	5	6.67	0.52	5
EN1	6.83	0.41	5	6.33	1.03	5	5.67	1.03	5
EN2	4.83	2.14	5	5.50	1.22	5	5.17	2.32	5
EN3	7.00	0.00	5	7.00	0.00	5	7.00	0.00	5

Table 3. Descriptive results from the experts' evaluation for each group of proposals and global assessment regarding the pertinence, relevance and clarity

	Pertinence			Relevance			Clarity		
	AVG	STD	N	AVG	STD	N	AVG	STD	N
TR	6.24	1.00	20	5.76	1.26	20	6.24	1.04	20
US/UX	6.48	1.47	20	6.62	1.32	20	6.24	1.51	20
EN	6.19	1.64	15	6.25	1.13	15	5.88	1.67	15
Global	6.27	1.30	55	6.22	1.30	55	6.11	1.42	55

Related to their responses about each proposal, as previously said, the expert had to assess the change proposal regarding the pertinence, relevance and clarity. Also in each question related to a proposal, the expert could introduce qualitative feedback through texting its opinion. Table 2 gathers the average mark, standard deviation and number of responses collected for each change/improvement proposal in terms of pertinence, relevance and clarity. Also, Table 3 gathers the same information but showing it in groupings related to the main topic associated to each change/improvement proposal (trust, usability / user experience and engagement) as well as the global average, standard deviation and number of responses collected in the assessment questionnaire.

The calculations and original responses retrieved from the experts can be checked in the sheet 2 of the following spreadsheet https://docs.google.com/spreadsheets/d/1dO72ZiHTt83UI2_cfjSd5sO1M109TXdO5rysCqgIp94/edit?usp=sharing.

In general, the average mark of the assessment in each question and grouping topic could be considered as good: most of the results are in the Q1 (score 5.5).

This Q1 score is not achieved in the proposed change EN2 (*inclusion of textual feedback related to user responses including information that may be relevant*) pertinence and clarity, TR1 (*modify the text and appearance from the invitation letter to the questionnaire*) regarding its relevance and US/UX3 proposed change (*deactivation of control elements when an action is initiated*).

Regarding the qualitative comments introduced by the experts in their feedback, the following could be highlighted:

- Comments with recommendations about visual design and layout as well as minor changes in the text of the proposed invitation letter and proposed introduction text to the questionnaire.
- Comments about the fact that many users will not know previously the OEEU's visual brand, so many graduates would not develop positive feelings regarding to trust in TR2 proposal.
- A comment regarding to US/UX3 (*deactivation of control elements when an action is initiated*) proposed change where the expert explains that he/she "is not aware about what implies this change".
- Very positive comments regarding the US/UX4 proposal (*in related elements, instead of having smaller and more specific groupings, use some larger grouping, following the Gestalt law on grouping*).
- Comments related to EN1 proposed improvement (in the questions related to the community in which they live, change the drop-down selector for a map with the autonomous communities of Spain) to include something similar for graduates that do not live now in Spain and live abroad (instead of selecting Spain autonomous communities, select countries, etc.).
- Two positive comments and another two expressing doubts about EN2 proposal (*inclusion of textual feedback related to user responses including information that may be relevant*). The positive comments explain that this change could lead to engage users by taking advantage of their curiosity. The other two explain that these extra screens and personalized feedback could break the users' trust in the data anonymization and introduce some distortions in the questionnaire.
- Some comments regarding little details that could improve the notifications. For example: the text to accept the reception of notifications should be "Yes, I accept" instead of "Ok, I accept" or introduce information about how much time will take to the user to complete the questionnaire if the he/she continues it.

In general terms, the feedback from the experts about the proposed changes/improvements for the questionnaire is very positive. Most part of the scores gathered by the experts are in the top first quartile of the scale (values 1-7), so can be accepted “as is” to be implemented in the questionnaire of course, after a final evaluation of the convenience with the project managers and OEEU coordinators -.

On the other hand, the experts raised some doubts in other elements or certain assessment points, like the pertinence and clarity of EN2 proposal (textual feedback), the relevance of TR1 proposal (modifying the text and appearance of the invitation letter) or the relevance of the US/UX3 proposal (deactivation of control elements when an action is initiated). In these cases, all the evaluations exceeded the Q2 score (4.0 value), so still they can be considered as well perceived changes but, in any case, these should be reviewed again by the researchers, in order to improve them or discard certain proposals if there is no possible improvement for that.

Despite some of these changes that are not fully supported by experts usually are backed by other authors in the literature, researchers should follow a pragmatic approach that ensures the right application of this kind of changes/improvements for the specific case of the OEEU’s questionnaire and its context.

Also, as previously commented, all these changes and improvement proposals will be validated again with the OEEU project coordinators and OEEU project managers before implementing them in the final version of the questionnaire that will be public in April 2017.

6 Conclusions

This paper presents a research focused on improving the success/completion ratio in large surveys. In this case, the large survey is the questionnaire produced by the Spanish Observatory for University Employability and Employment and that will be publicly available for graduates of master degree in April 2017. This questionnaire is composed by about 32 and 60 questions and between 86 and 181 variables to be measured. The research is based on the previous experience of a past questionnaire proposed also by the Observatory composed also by a large amount of questions and variables to be measured.

Analyzing the target population of the questionnaire (also comparing with the target population of the previous questionnaire) and reviewing the literature, the researchers have designed 11 proposals for changes related to the questionnaire that could improve the users’ completion and success ratios (changes that could improve the users’ trust in the questionnaire, the questionnaire usability and user experience or the users’ engagement to the questionnaire). These changes are planned to be applied in the questionnaire in two main different experiments based on A/B test methodologies that will allow researchers to measure the effect of the changes in different populations and in an incremental way.

The proposed changes have been assessed by five experts through an evaluation questionnaire. In this questionnaire, researchers gathered the score of each expert regarding to the pertinence, relevance and clarity of each change proposed. Regarding the results of this evaluation questionnaire, the reviewers fully supported 8 out of the 11 changes proposals, so they could be introduced in the questionnaire with no variation. On the other hand, 3 of the proposed changes or improvements are not fully supported by the experts (they have not received a score in the top first quartile of the 1-7 Likert scale). These changes will not be discarded immediately, because despite they have not received a Q1 score, they received a score within the second quartile. Instead of being discarded, these changes will be reviewed again by the researchers and the Observatory staff in order to adequate them to the questionnaire. If there is no possibility to adequate them to the OEEU's questionnaire context, finally they will be finally rejected.

After all this work, research and validation processes, the future work is to implement all the accepted changes and variations in the OEEU's questionnaire for graduates of master studies and study what of these changes lead to a real improvement in the completion/success ratio related to the questionnaire.

Acknowledgments

The research leading to these results has received funding from "la Caixa" Foundation. Also, the author Juan Cruz-Benito would like to thank the European Social Fund and the *Consejería de Educación* of the *Junta de Castilla y León* (Spain) for funding his predoctoral fellow contract. This work has been partially funded by the Spanish Government Ministry of Economy and Competitiveness throughout the DEFINES project (Ref. TIN2016-80172-R).

References

1. Jarrett, C., Gaffney, G.: Forms that work: Designing Web forms for usability. Morgan Kaufmann (2009)
2. Dillman, D.A.: Mail and internet surveys: The tailored design method. Wiley New York (2000)
3. Michavila, F., Martín-González, M., Martínez, J.M., García-Peñalvo, F.J., Cruz-Benito, J.: Analyzing the employability and employment factors of graduate students in Spain: The OEEU Information System. Third International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'15). ACM Inc., Porto, Portugal (2015)
4. Michavila, F., Martínez, J.M., Martín-González, M., García-Peñalvo, F.J., Cruz-Benito, J.: Barómetro de Empleabilidad y Empleo de los Universitarios en España, 2015 (Primer informe de resultados). (2016)
5. Cruz-Benito, J., Therón, R., García-Peñalvo, F.J., Martín-González, M.: Herramienta para la validación de elementos de mejora UX/Engagement para los cuestionarios de recogida de información de egresados en el contexto del Observatorio de Empleabilidad y Empleo Universitarios (OEEU). GRIAL Research Group. University of Salamanca (2017)
6. Anderson, S.P.: Seductive Interaction Design: Creating Playful, Fun, and Effective User Experiences, Portable Document. Pearson Education (2011)
7. Kveton, P., Jelínek, M., Klimusová, H., Voboril, D.: Data collection on the Internet: Evaluation of Web-based questionnaires. Studia psychologica 49, 81 (2007)

8. Flavián, C., Gurrea, R., Orús, C.: The effect of product presentation mode on the perceived content and continent quality of web sites. Online Information Review 33, 1103-1128 (2009)
9. Huber, W., Vitouch, P.: Usability and accessibility on the internet: Effects of accessible web design on usability. In: International Conference on Computers for Handicapped Persons, pp. 482-489. Springer, (2008)
10. Seckler, M., Tuch, A.N., Opwis, K., Bargas-Avila, J.A.: User-friendly locations of error messages in web forms: Put them on the right side of the erroneous input field. Interacting with Computers 24, 107-118 (2012)
11. Shneiderman, B.: Designing trust into online experiences. Communications of the ACM 43, 57-59 (2000)
12. Recabarren, M., Nussbaum, M.: Exploring the feasibility of web form adaptation to users' cultural dimension scores. User Modeling and User-Adapted Interaction 20, 87-108 (2010)
13. Pengnate, S.F., Sarathy, R.: An experimental investigation of the influence of website emotional design features on trust in unfamiliar online vendors. Computers in Human Behavior 67, 49-60 (2017)
14. Sayago, S., Blat, J.: Some Aspects of Designing Accessible Online Forms for the Young Elderly. In: WEBIST (2), pp. 13-17. (2007)
15. Stieger, S., Reips, U.-D.: What are participants doing while filling in an online questionnaire: A paradata collection tool and an empirical study. Computers in Human Behavior 26, 1488-1495 (2010)
16. ISO/DIS: Draft BS ENISO 9241-220 Ergonomics of human-computer interaction. Part 220: Processes for enabling, executing and assessing human-centred design within organizations. (2016)
17. ISO/DIS: Ergonomics of human-system interaction — Part 125: Guidance on visual presentation of Information. (2013)
18. Bevan, N., Carter, J., Harker, S.: ISO 9241-11 revised: What have we learnt about usability since 1998? In: International Conference on Human-Computer Interaction, pp. 143-151. Springer, (2015)
19. Bargas-Avila, J., Brenzikofer, O.: Simple but crucial user interfaces in the world wide web: introducing 20 guidelines for usable web form design (2011)
20. Bevan, N., Spinhof, L.: Are guidelines and standards for web usability comprehensive? In: International Conference on Human-Computer Interaction, pp. 407-419. Springer, (2007)
21. Seckler, M., Heinz, S., Forde, S., Tuch, A.N., Opwis, K.: Trust and distrust on the web: User experiences and website characteristics. Computers in Human Behavior 45, 39-50 (2015)
22. Bargas-Avila, J.A., Brenzikofer, O., Tuch, A.N., Roth, S.P., Opwis, K.: Working towards usable forms on the worldwide web: optimizing multiple selection interface elements. Advances in Human-Computer Interaction 2011, 4 (2011)
23. Couper, M.P., Tourangeau, R., Conrad, F.G., Crawford, S.D.: What they see is what we get response options for web surveys. Social science computer review 22, 111-127 (2004)
24. Sánchez-Prieto, J.C., Olmos-Migueláñez, S., García-Peñalvo, F.J.: Informal tools in formal contexts: Development of a model to assess the acceptance of mobile technologies among teachers. Computers in Human Behavior 55, 519-528 (2016)

6.9 Appendix I. Herramienta para la validación de elementos de mejora UX/Engagement para los cuestionarios de recogida de información de egresados en el contexto del Observatorio de Empleabilidad y Empleo Universitarios (OEEU)

Herramienta para la validación de elementos de mejora UX/Engagement para los cuestionarios de recogida de información de egresados en el contexto del Observatorio de Empleabilidad y Empleo Universitarios (OEEU)

Elaborado por: Juan Cruz-Benito, Roberto Therón,
Francisco J. García-Peñalvo, Martín Martín-González
DOI: http://doi.org/10.5281/zenodo.322575

Validado por:
Fecha de validación:

Autor/es: Juan Cruz-Benito, Roberto Therón,
Francisco J. García-Peñalvo, Martín Martín-González

Introducción

La recogida de información mediante cuestionarios y entrevistas es uno de los métodos más conocidos y utilizados en la actualidad, tanto en el medio físico como en el digital.
Es común que cualquier web que se visite cuente con algún formulario para la entrada de información, ya sea como punto de contacto, como parte del acceso a partes privadas del sistema, como parte de un proceso de pago, etc. Los formularios están tan integrados en la interacción del usuario dentro de una web, que puntualmente se relativiza su importancia y se presupone que el usuario va a completarlo por el mero hecho de utilizarlos habitualmente. Sin embargo, esto no es así.

Debido a que los usuarios utilizan los formularios habitualmente, en los últimos años estos usuarios han desarrollado ciertas tendencias y conductas frente a dichas herramientas de entrada de información. Por ejemplo, se ha comprobado que los usuarios [1]:

- Confían más en los sitios web, incluso estando más dispuestos a realizar acciones complejas (a todos los niveles), como compras, pagos, etc.
- Protegen más su información, están menos dispuestos a revelar información personal
- Demandan mejores productos, son menos tolerantes a malos formularios

Durante los últimos años se ha realizado multitud de trabajo en relación a los cuestionarios, detectándose que los usuarios tienen cierta reticencia hacia completar un formulario desde antes de rellenarlo [1]. Esto impone ciertos problemas respecto a la consecución de objetivos de recogida de información intrínsecos a cualquier formulario.

Respecto a los tipos de usuarios que completan formularios, se pueden establecer distintos perfiles [1]:

1. Usuarios que leen: aquellos que leen el formulario cuidadosamente
2. Usuarios apresurados: completan lo más rápido que pueden los campos de los formularios, leyendo únicamente el formulario cuando creen que es necesario.
3. Usuarios que se niegan: son aquellos que creen que no tienen que hacer nada con el formulario.

Además del perfil de usuario, hay diversos factores relevantes que influyen decisivamente en la intención de completar un cuestionario por parte de usuarios [2]:

- Confianza
- Costes (sociales, individuales). Estos costes se pueden vislumbrar en los sentimientos negativos del usuario frente al cuestionario, en la incomodidad, etc.
- Recompensas

Según la literatura, e íntimamente relacionado con la Teoría de Intercambio Social [2], algunos autores [1] distinguen tres capas en los formularios: relación, conversación y apariencia.

1. La relación de un formulario se basa en la relación que tiene quién hace las preguntas con quién las responde.

2. La conversación de un formulario va desde las cuestiones que se preguntan, a las instrucciones que se dan o la organización de las preguntas según su temática.
3. La apariencia del formulario es la imagen que proyecta: colocación del texto, gráficos, áreas de entrada de datos, color, etc.

Mejorar estos factores la relación del usuario, es más fácil que éste esté dispuesto a participar y completar su tarea dentro del cuestionario. Y son en ellas donde se incidirá, partiendo de los puntos de vista de la usabilidad y el *engagement* (enganche), como se comentará más adelante.

1.1 El contexto: Observatorio de Empleabilidad y Empleo Universitarios (OEEU)

Durante los meses de Junio – Julio de 2015 el Observatorio de Empleabilidad y Empleo Universitarios (OEEU) contactó a varios miles de egresados universitarios españoles (133588) – a través de las universidades (48, públicas y privadas) donde cursaron estudios de grado durante el curso 2009-2010 - para invitarles a rellenar un cuestionario [3, 4].
Este cuestionario contaba con una parte común con 60 preguntas y 167 variables medidas, además de 3 itinerarios específicos (ramificaciones) en función de las respuestas de los usuarios. El primer itinerario añade 3 preguntas y 9 variables medidas más. El segundo, añade 24 preguntas y 70 variables. Por último, el tercer itinerario añade 32 cuestiones más y 112 variables. Por tanto el cuestionario varía entre 63-92 cuestiones y 176-279 variables en función del itinerario que siga el usuario. Se puede afirmar sin lugar a dudas, que el cuestionario es muy extenso.

La cantidad de usuarios que iniciaron el cuestionario fueron 13006 (9.74% del total de la población), de los cuales lo finalizaron 9617 (7.20% del total de la población, 73.94% del total de cuestionarios iniciados).
Los datos descriptivos en cuanto a la edad de los participantes en el cuestionario fueron los siguientes:

```
count    12109.000000
mean        32.525972
std          7.018282
min         25.000000
25%         28.000000
50%         30.000000
75%         34.000000
max         80.000000
```

En cuanto al sexo, el 56,05% (7290) de los usuarios que respondieron el cuestionario fueron mujeres y el 43,94% (5716) hombres. En relación a la nacionalidad, el 98,54% (11672) de los usuarios son españoles y 1,46% (173) extranjeros.

Sobre los usuarios que abandonaron el cuestionario, los cuartiles de tasa de abandono en función de la pantalla del cuestionario donde lo dejaron fue:

```
count    3389.000000
25%         4.000000
50%         5.000000
75%         7.000000
```

Es decir, el 25% de los usuarios abandonaron en la pantalla 4 o antes, otro 25% abandonaron entre las pantallas 4 y 5 del cuestionario, otro 25% entre las pantallas 5 y 7 y otro 25% entre la pantalla 7 y el final (dependiendo de su itinerario).

Ahora, en 2017, se va a volver a realizar un proceso de recogida de información similar al realizado durante 2015. En este caso se va a recoger información sobre egresados de estudios de máster durante el curso 2013-2014. Para ello, se ha propuesto un cuestionario compuesto por entre 32 y 60 preguntas (el cuestionario cuenta con varios itinerarios en función de las respuestas del usuario) y entre 86 y 181 variables medidas. En principio, se puede considerar que a pesar de las diferencias, es un cuestionario grande y comparte parte de la problemática del anterior en cuanto a las dificultades o retos que puede presentar para ser completado por los usuarios.

Actualmente, a día 14 de Febrero, se cuenta con datos de 28744 personas provenientes de 32 universidades españolas públicas y privadas. De estas personas se tienen los siguientes datos:
Datos descriptivos en cuanto a la edad de la población a la que se dirigirá el cuestionario:

```
count    28744.000000
mean        35.854370
std         15.852381
min          5.000000
25%         28.000000
50%         31.000000
75%         38.000000
max        117.000000
```

En los datos de la edad, evidentemente se nota el envejecimiento de la población respecto a la del cuestionario anterior. Esto es normal pensando en que la edad de entrada a un máster es más alta que la que se tiene al acceder a un grado (al menos de forma habitual).

Sobre el sexo de la población a la que se dirigirá el cuestionario, el 55,2% (16385) son mujeres y el 44,8% (13317) son hombres. En relación a la nacionalidad, este es el aspecto en el que más se diferencia la población actual (egresados de máster) de la del estudio de egresados de grado. En esta ocasión, la proporción de estudiantes extranjeros es mayor, habiendo un 88,11% (25318) de estudiantes de nacionalidad española frente al 11,88% (3414) estudiantes con nacionalidad extranjera.

En líneas generales, se podría decir que las poblaciones (poniendo en contexto cada una de ellas) no son muy distintas. En este sentido, se podría destacar la diferencia principal existente en cuanto a la nacionalidad. Esta diferencia podría llevar a considerar tratar de forma distinta algunos aspectos del cuestionario para ajustarse a posibles diferencias culturales. En este caso, no va a haber distinción cultural a la hora de diseñar, presentar o realizar el cuestionario. Esto podría considerarse una limitación del estudio en relación a los objetivos que se plantearán a continuación.

1.2 Objetivos

Los objetivos principales del experimento que se está diseñando, y que se presentará a continuación, son:

- Estudiar cómo mejorar el ratio de inicio del cuestionario (anteriormente, cercanos a un 9,7% del total de la población).
- Estudiar cómo mejorar el ratio de finalización del cuestionario (anteriormente 73,94% de tasa de finalización).

Además de estos objetivos fundamentales, se podría proponer otro objetivo relacionado con el segundo, y es que en caso de abandono, los usuarios hayan completado todas las pantallas posibles (obteniéndose más información aunque abandonen).

1.3 Propuesta

Para la nueva versión del cuestionario, se considera que se pueden mejorar diversos puntos en relación al cuestionario implementado en 2015 y a las formas de aumentar la participación por parte de los usuarios.

Para implementar estas mejoras, se propone la realización de dos experimentos en paralelo:

- Estudio sobre cómo mejorar los procesos de invitación a los egresados.
- Estudio sobre qué mejoras se pueden implementar en los cuestionarios para mejorar los ratios de participación y finalización.

A continuación se comentarán los aspectos clave de cada uno de los estudios, indicando también los cambios principales que se implementarían, etc. Posteriormente en el cuestionario para expertos, el lector podrá opinar sobre dichos elementos.
Este cuestionario para expertos se basa en la propuesta de Sánchez-Prieto, et. al. [5] y en ella se evalúa mediante una escala Likert 1-7 la pertinencia del cambio propuesto, la claridad del mismo y su importancia, además de permitir un comentario por parte del experto sobre cualquier cuestión relacionada.

1.3.1 Estudio sobre cómo mejorar los procesos de invitación a los egresados

En el caso de los cuestionarios producidos por el OEEU, es necesario tener en cuenta un factor fundamental: la privacidad del usuario prima por encima de todo (entre otros motivos, debido a los datos sensibles que se manejan). Este proyecto es absolutamente respetuoso y cumple con la Ley Orgánica de Protección de Datos (LOPD), teniendo registrada su base de datos incluso ante las autoridades españolas como salvaguarda de los mismos datos.

Debido a las restricciones de privacidad impuestas desde el proyecto, el Observatorio no guarda datos que permitan relacionar a una persona con su información. Es decir, no se relaciona información con nombres, DNI, fecha exacta de nacimiento, etc. La única excepción es que se le ofrece a los usuarios la opción de incluir su correo electrónico para informarles de las investigaciones del Observatorio o los resultados del sorteo de unas *tablets* celebrado entre los egresados que completan el cuestionario.

Atendiendo a estas restricciones, y debido a que el email si se llega a obtener es al final de todo el proceso, son las universidades las que contactan a sus egresados ofreciéndoles participar en el proceso de los cuestionarios, indicándoles a su vez que hay un sorteo entre los que completan dicho formulario y dándole el link (URL) personal a cada alumno para que realice el formulario.

A continuación, se presenta el correo electrónico base proporcionado por el Observatorio de Empleabilidad y Empleo Universitario a las universidades. Éstas pudieron utilizarlo o no, siendo responsabilidad de cada una de ellas su uso y modificación.

> Estimado/a egresado/a (si la Universidad quiere, puede enviar el correo personalizado indicando el nombre del/de la egresado/a):
>
> Desde la Universidad (nombre de la Universidad) te pedimos que dediques 15 minutos de tu tiempo a responder un cuestionario —que por primera vez se aplica a todas las universidades españolas—, para conocer la situación laboral de los titulados universitarios.
>
> Los datos introducidos serán tratados de manera anónima por la Universidad (nombre de la Universidad) y el Observatorio de Empleabilidad y Empleo Universitarios (OEEU), cumpliendo con la Ley Orgánica de Protección de Datos de Carácter Personal (LOPD). Los resultados agregados podrás consultarlos próximamente en la página web del Observatorio (www.oeeu.org).
>
> Entre los que respondan la encuesta se sortearán, durante el mes de junio de 2015, 10 tabletas Samsung Galaxy Tab 4.
>
> Pincha aquí para realizar la encuesta: «URL»
>
> ¡Muchas gracias por colaborar!
>
> Un cordial saludo,
>
> Nombre de la Universidad
> Correo electrónico de contacto
> Teléfono de contacto
>
> Observatorio de Empleabilidad y Empleo Universitarios
> udyc@oeeu.org
> 913364185

La propuesta de experimento en cuanto a este aspecto se basa en, utilizando varias universidades participantes de la edición anterior (por ejemplo: Universidad de Salamanca, Universidad Politécnica de Madrid, etc.) que usasen este texto de invitación, enviar una versión distinta de la invitación (modificando el tono y contenido textual del mensaje, más aportando otro diseño del mismo [6]) para probar cómo cambian los ratios de apertura e inicio de participación del cuestionario respecto a la edición anterior. Del mismo modo, en esta edición, se propone que en esta edición dichas universidades envíen esa versión modificada, mientras que el resto sigan enviando una versión simplemente actualizada (con los datos de esta edición) de la invitación anterior. Con esta propuesta se permite ver el efecto de los cambios en la invitación (mediante el uso de metodologías de *test A/B*) teniendo en cuenta varias cosas:

- El contexto de cada universidad participante es distinto (factores poblacionales, económicos, etc.). Por ello se hacen cambios específicos para universidades participantes en ambas convocatorias de recogida de datos.
- La población del estudio ha cambiado de la primera edición de recogida de datos a la segunda (edad, formación, etc.). Por ello se validan los cambios entre diversas universidades dentro de la misma edición de la recogida de datos.

En la sección de validación por expertos que hay en este documento, se comentarán los cambios que se van a introducir. En cualquier caso, los cambios a introducir vienen limitados por los diversos condicionantes del proyecto relacionados con la privacidad (no es posible utilizar plataformas de envío de correo externas, etc.) y se enfocan fundamentalmente en las cuestiones de mejorar la confianza y relación entre el usuario y la entidad que propone el cuestionario (OEEU).

1.3.2 Estudio sobre qué mejoras se pueden implementar en los cuestionarios para mejorar los ratios de participación y finalización

En cuanto a la parte del estudio relacionada con los cambios en el cuestionario, se proponen diversas modificaciones en varios niveles [7].

El planteamiento general del estudio es realizar un test A/B de tres variantes (A/B/C). Se propone una variante base (A) que siga las líneas generales del cuestionario de la anterior edición (disponible en http://gredos.usal.es/jspui/bitstream/10366/127374/5/Anexos_OEEU_2015.pdf), del cual se tiene cierta idea de la eficiencia, etc., junto a otras dos variantes (B y C) que cambian determinados aspectos relacionados con la Teoría del Intercambio Social [2].

En general, la variante B del test (experimento) se refiere a cambios relacionados con la relación del participante (que contesta) y de quién propone el cuestionario (primera capa de la teoría) junto a cambios relacionados con la apariencia (tercera capa de la teoría) [8, 9]. De forma más amplia, esta variante B se basa en la confianza entre las partes [10, 11], más mejoras y cambios respecto a la experiencia de usuario [12], usabilidad [9] y diseño de la interfaz del cuestionario [13, 14] .

Por su lado, la variante C del test incluye los cambios propuestos en la variante B más otros cambios relacionados con la relación entre las partes del cuestionario (primera capa) y la conversación entre ellas (segunda capa). Desde este punto de vista, la variante C se centrará en cuestiones relacionadas con el *engagement* de los usuarios [6].

La mayoría de las propuestas que se presentan se encuentran sugeridas y soportadas por literatura previa. En el caso concreto las propuestas relacionadas con la usabilidad y los cambios en el diseño, se siguen recomendaciones de autores, normas ISO de usabilidad y recomendaciones y las guías HHSI [15-17].

En cualquier caso, los tres elementos dispondrán de ciertas recompensas que ya se ofrecieron en el anterior proceso de recogida de datos. Por ejemplo, en esta ocasión volverá a haber un sorteo de dispositivos electrónicos entre aquellas personas que completan el cuestionario. También se mantendrá la comunicación con aquellos usuarios que quieran recibir las últimas novedades de la actividad del Observatorio.

Respecto a algunas de las componentes como la edad, la discapacidad, u otras situaciones y contextos personales de los usuarios, en este caso se van a obviar (exceptuando normas de accesibilidad generales) debido a que el experimento no se va a centrar en aspectos concretos relacionados con posibles subgrupos dentro de la población del estudio [9, 18]. Se asume que esto constituye una limitación del estudio que se va a hacer.

El efecto de los cambios se va a medir de dos formas:

- Comprobando los datos en cuanto al acceso a los cuestionarios, la compleción de cada parte del cuestionario y la compleción final del mismo (pantallas completadas, puntos de abandono, etc),
- Evaluando los *paradata* [19]. Los *paradata* de un cuestionario son los datos auxiliares que describen un proceso, como los tiempos de respuesta, clics, procesos de scroll, etc. En este caso, los *paradata* tendrán que ver con los tiempos que tarda el usuario en completar la tarea de responder cada página de preguntas, el tiempo en completar el cuestionario en total, los accesos al cuestionario personalizado, etc. Estos *paradata* no podrán compararse con datos similares de la ronda anterior de recolección de datos en grado, ya que no se hizo nada similar.

Habitualmente, en este tipo de investigación, se suele hacer un cuestionario a los usuarios sobre su opinión acerca de cómo se han sentido frente al cuestionario, cómo han podido resolver la tarea, etc. En este caso, debido a la extensión del cuestionario que han de rellenar y del carácter del proyecto, no se va a realizar esta investigación de ese modo. Esto supone una limitación en cuanto a la riqueza de los resultados que se pueden obtener. Lo que sí se va a plantear, es que los alumnos que decidan dar voluntariamente su correo electrónico podrán ser preguntados en un nuevo cuestionario especializado acerca de estas cuestiones.

En la tabla 1 se presentan las propuestas de cambio y su dimensión de actuación en cuanto a las capas de la Teoría del Intercambio Social y a las áreas de la Interacción Persona-Ordenador que tratan.

Estas medidas propuestas se detallan con ejemplos concretos e imágenes (en los casos que se requiera) en el apartado del cuestionario de validación propiamente dicho.

Tabla 1. Propuestas de cambio en las versiones B y C de los cuestionarios respecto a la Teoría de Intercambio Social y elementos/áreas de la Interacción Persona-Ordenador sobre los que se actuará.

Capa de la Teoría de Intercambio Social	Relación	Conversación	Apariencia	
Dimensión de actuación	Confianza	*Engagement*	Usabilidad / Experiencia de usuario	Diseño
Adecuación de la imagen al resto de los productos del observatorio	X		X	X
Inclusión del logotipo del observatorio + logotipo de la universidad	X			
Inclusión de barra de progreso			X	X
Foco visual sobre la acción concreta			X	X
Desactivación de elementos de control una vez iniciada una acción (ejemplo: desactivar los botones una vez que se pulsa uno de ellos hasta que se termina de realizar la acción deseada).	X		X	
En elementos relacionados, en lugar de tener agrupaciones más pequeñas y específicas, utilizar alguna más grande siguiendo la ley de la Gestalt sobre agrupación. En este caso la cabecera de la tabla quedaría fijada mientras que en el contenido se pueda hacer *scroll*.			X	X
Cambios en el texto de introducción al cuestionario	X	X		
En las preguntas relacionadas con la comunidad en la que viven, cambiar un selector *drop-down* por un mapa con las comunidades autónomas de España en la que el usuario pueda pinchar en la correspondiente		X		X
Inclusión de *feedback* textual relacionado con las respuestas del usuario incluyendo información que pueda ser relevante.		X		X
Inclusión de notificaciones *push* web que permitan mandar al usuario mensajes de ánimo si abandona el cuestionario		X		

Agradecimientos

La investigación que ha dado lugar a estos resultados ha sido impulsada por la Obra Social “la Caixa”. Del mismo modo, el autor Juan Cruz-Benito agradece a la Junta de Castilla y León y al Fondo Social Europeo la financiación de su contrato de investigador predoctoral.

Referencias

[1] C. Jarrett and G. Gaffney, *Forms that work: Designing Web forms for usability*. Morgan Kaufmann, 2009.
[2] D. A. Dillman, *Mail and internet surveys: The tailored design method*. Wiley New York, 2000.
[3] F. Michavila, M. Martín-González, J. M. Martínez, F. J. García-Peñalvo, and J. Cruz-Benito, "Analyzing the employability and employment factors of graduate students in Spain: The OEEU Information System," presented at the Third International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'15), Porto, Portugal, 2015.
[4] F. Michavila, J. M. Martínez, M. Martín-González, F. J. García-Peñalvo, and J. Cruz-Benito, "Barómetro de Empleabilidad y Empleo de los Universitarios en España, 2015 (Primer informe de resultados)," 2016.
[5] J. C. Sánchez-Prieto, S. Olmos-Migueláñez, and F. J. García-Peñalvo, "Informal tools in formal contexts: Development of a model to assess the acceptance of mobile technologies among teachers," *Computers in Human Behavior,* vol. 55, pp. 519-528, 2016.
[6] S. P. Anderson, *Seductive Interaction Design: Creating Playful, Fun, and Effective User Experiences, Portable Document*. Pearson Education, 2011.
[7] P. Kveton, M. Jelínek, H. Klimusová, and D. Voboril, "Data collection on the Internet: Evaluation of Web-based questionnaires," *Studia psychologica,* vol. 49, no. 1, p. 81, 2007.
[8] C. Flavián, R. Gurrea, and C. Orús, "The effect of product presentation mode on the perceived content and continent quality of web sites," *Online Information Review,* vol. 33, no. 6, pp. 1103-1128, 2009.
[9] W. Huber and P. Vitouch, "Usability and accessibility on the internet: Effects of accessible web design on usability," in *International Conference on Computers for Handicapped Persons*, 2008, pp. 482-489: Springer.
[10] M. Seckler, S. Heinz, S. Forde, A. N. Tuch, and K. Opwis, "Trust and distrust on the web: User experiences and website characteristics," *Computers in Human Behavior,* vol. 45, pp. 39-50, 2015.
[11] B. Shneiderman, "Designing trust into online experiences," *Communications of the ACM,* vol. 43, no. 12, pp. 57-59, 2000.
[12] M. Recabarren and M. Nussbaum, "Exploring the feasibility of web form adaptation to users' cultural dimension scores," *User Modeling and User-Adapted Interaction,* vol. 20, no. 1, pp. 87-108, 2010.
[13] S. F. Pengnate and R. Sarathy, "An experimental investigation of the influence of website emotional design features on trust in unfamiliar online vendors," *Computers in Human Behavior,* vol. 67, pp. 49-60, 2017.
[14] M. Seckler, A. N. Tuch, K. Opwis, and J. A. Bargas-Avila, "User-friendly locations of error messages in web forms: Put them on the right side of the erroneous input field," *Interacting with Computers,* vol. 24, no. 3, pp. 107-118, 2012.
[15] ISO/DIS, "Draft BS ENISO 9241-220 Ergonomics of human-computer interaction. Part 220: Processes for enabling, executing and assessing human-centred design within organizations," 2016.

[16] ISO/DIS, "Ergonomics of human-system interaction — Part 125: Guidance on visual presentation of Information," 2013.

[17] N. Bevan, J. Carter, and S. Harker, "ISO 9241-11 revised: What have we learnt about usability since 1998?," in *International Conference on Human-Computer Interaction*, 2015, pp. 143-151: Springer.

[18] S. Sayago and J. Blat, "Some Aspects of Designing Accessible Online Forms for the Young Elderly," in *WEBIST (2)*, 2007, pp. 13-17.

[19] S. Stieger and U.-D. Reips, "What are participants doing while filling in an online questionnaire: A paradata collection tool and an empirical study," *Computers in Human Behavior,* vol. 26, no. 6, pp. 1488-1495, 2010.

Evaluación de los cambios propuestos para la mejora de elementos UX/Engagement para los cuestionarios de recogida de información de egresados en el contexto del Observatorio de Empleabilidad y Empleo Universitarios (OEEU)

A continuación se expondrán los distintos cambios propuestos para cada experimento y sub-experimento comentados anteriormente. Podrás opinar acerca de su pertinencia, claridad e importancia y hacer un comentario abierto en el que comentar tu opinión de forma no estructurada.

Si lo prefieres, hay una versión online del cuestionario que puedes completar en la URL https://goo.gl/forms/kXXwQQcV6VrWAjyF2

Gracias por tu participación.

Sexo:

□ Hombre
□ Mujer

Edad:

□ Hasta 30 años
□ Entre 31 y 40 años
□ Entre 41 y 50 años
□ Entre 51 y 60 años
□ Más de 60 años

Área de conocimiento:

□ Artes y Humanidades
□ Ciencias
□ Ciencias de la Salud
□ Ciencias Sociales y Jurídicas
□ Ingeniería y arquitectura

¿Cuál es tu campo de especialización (informática, economía, política, etc.)? (Respuesta abierta)

Propuesta para el texto de invitación a los cuestionarios

El texto original de la invitación aparece a continuación. En principio a la mayoría de universidades se les ofrecerá utilizarlo del mismo modo que en la anterior edición. Este texto adjunto simplemente cambiaría indicando que en esta edición se trata de un cuestionario relacionado con estudios de máster y haciendo referencia a la edición anterior de grado.

> Estimado/a egresado/a (si la Universidad quiere, puede enviar el correo personalizado indicando el nombre del/de la egresado/a):
>
> Desde la Universidad (nombre de la Universidad) te pedimos que dediques 15 minutos de tu tiempo a responder un cuestionario —que por primera vez se aplica a todas las universidades españolas—, para conocer la situación laboral de los titulados universitarios.
>
> Los datos introducidos serán tratados de manera anónima por la Universidad (nombre de la Universidad) y el Observatorio de Empleabilidad y Empleo Universitarios (OEEU), cumpliendo con la Ley Orgánica de Protección de Datos de Carácter Personal (LOPD). Los resultados agregados podrás consultarlos próximamente en la página web del Observatorio (www.oeeu.org).
>
> Entre los que respondan la encuesta se sortearán, durante el mes de junio de 2015, 10 tabletas Samsung Galaxy Tab 4.
>
> Pincha aquí para realizar la encuesta: «URL»
>
> ¡Muchas gracias por colaborar!
>
> Un cordial saludo,
>
> Nombre de la Universidad
> Correo electrónico de contacto
> Teléfono de contacto
>
> Observatorio de Empleabilidad y Empleo Universitarios
> udyc@oeeu.org
> 913364185

Por otra parte, a ciertas universidades participantes en la anterior edición (USAL, UPM, etc.) se les ofrecerá pasar una invitación al cuestionario similar a la siguiente. Entre los cambios propuestos, destaca la inclusión del logotipo de la Universidad que envía la invitación, la inclusión del logotipo del Observatorio de Empleabilidad y Empleo Universitario, un diseño acorde a los colores y tipografías que se utilizarán en el cuestionario, y un cambio en el texto de forma que sea una invitación más personal hacia el egresado.

Estos cambios tienen como intención mejorar la confianza del usuario hacia el cuestionario y la actividad del Observatorio.

VNIVERSIDAD D SALAMANCA
CAMPUS OF INTERNATIONAL EXCELLENCE

Estimado/a egresado/a (si la Universidad quiere, puede enviar el correo personalizado indicando el nombre del/de la egresado/a):

Desde la Universidad (nombre de la Universidad) y el Observatorio de Empleabilidad y Empleo Universitarios (OEEU, http://oeeu.org) te invitamos a participar en un estudio para conocer la situación laboral de los titulados de máster en España. Para ello, te pedimos que dediques poco más de 15 minutos de tu tiempo en completar el cuestionario que te proponemos al final de este correo electrónico. En el cuestionario se te preguntará por diversos aspectos de tu vida laboral, personal (datos demográficos) y académica para intentar analizar cómo consiguen empleo los egresados de máster en España y qué aspectos afectan a su empleabilidad. Los datos que introduzcas serán tratados de manera absolutamente anónima y cumpliendo siempre con la Ley Orgánica de Protección de Datos de Carácter Personal (LOPD).

Como contraprestación a tu colaboración, sortearemos durante el mes de Julio de 2017, 10 tabletas Samsung Galaxy Tab 4. También serás el primero en conocer los resultados de aquellos estudios realizados el cuestionario.

Pincha aquí para comenzar con la encuesta: <<URL personalizada>>
¡Muchas gracias por tu colaboración!

Un cordial saludo,
Nombre de la Universidad
Correo electrónico de contacto
Teléfono de contacto

Observatorio de Empleabilidad y Empleo Universitarios
udyc@oeeu.org
913364185

Valoración de los cambios propuestos:

Pertinencia							Claridad							Importancia						
1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Observaciones

Propuesta de modificación del cuestionario para la variante B

Adecuación de la imagen (interfaz del cuestionario) al resto de los productos del observatorio. Modificación de colores, logotipos, tipografía, etc., de acuerdo a otros productos del Observatorio como la web http://datos.oeeu.org.

Pertinencia							Claridad							Importancia						
1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Observaciones

Inclusión del logotipo del Observatorio y el logotipo de la universidad del egresado en la cabecera de cada página/pantalla del cuestionario.

Pertinencia							Claridad							Importancia						
1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Observaciones

Inclusión de barra de progreso en las páginas/pantallas del cuestionario, de modo que el usuario sea capaz de conocer su avance en la tarea de completar el cuestionario y cuánto le queda para finalizarlo.

Pertinencia							Claridad							Importancia						
1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Observaciones

Foco visual sobre la acción concreta. En este caso, la web proporcionará un efecto de foco al usuario en el que percibirá siempre en el centro de la pantalla la tarea que tiene que resolver (típicamente contestar una pregunta del cuestionario), haciendo un efecto de desenfocado en los elementos que no sean fundamentales durante la resolución de la misma.

Para una obtener aclaración visual sobre esta propuesta, se pueden visitar los siguientes enlaces: https://drive.google.com/file/d/0BwS7cZg3riXtajJtNGhkMnlzXzg/view?usp=sharing, https://drive.google.com/file/d/0BwS7cZg3riXtWGk1bmlvSVB5dDg/view?usp=sharing

Pertinencia							Claridad							Importancia						
1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Observaciones

Desactivación de elementos de control una vez iniciada una acción (ejemplo: desactivar los botones una vez que se pulsa uno de ellos hasta que se termina de realizar la acción deseada). Con ello se busca que el usuario confíe la robustez del sistema y mejorar la usabilidad del mismo.

Para una obtener aclaración visual sobre esta propuesta, se puede visitar el siguiente enlace:
https://drive.google.com/file/d/0BwS7cZg3riXtVDY2dU9WRDlYWGs/view?usp=sharing

Pertinencia							Claridad							Importancia						
1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Observaciones

En elementos relacionados, en lugar de tener agrupaciones más pequeñas, utilizar alguna más grande siguiendo la ley de la Gestalt sobre agrupación. En este caso la cabecera de la tabla quedaría fijada mientras que en el contenido se pueda hacer *scroll*. En ella se busca que las grandes dimensiones de análisis en algunos puntos del cuestionario (competencias, métodos de acceso a un determinado empleo, etc.) se agrupen intentando evitar la fatiga del usuario en el uso de grandes tablas o elementos visuales complejos.

Para obtener una aclaración visual sobre este aspecto, se puede visitar el siguiente link https://drive.google.com/file/d/0BwS7cZg3riXtdmZqQzBHZXJVcmM/view?usp=sharing

Pertinencia							Claridad							Importancia						
1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Observaciones

Propuesta de modificación del cuestionario para la variante C

Cambios en el texto de introducción al cuestionario. En este caso se buscará un cambio en el texto de forma similar a la modificación en la invitación a los usuarios de un modo más personal y aportando argumentos de peso para influenciar una mejor percepción de lo que se va a realizar, mejorando la confianza en el cuestionario y la entidad que lo propone.

Texto de la anterior edición fue (en la variante A se actualizarán únicamente los datos del sorteo, etc.):

Autor/es: Juan Cruz-Benito, Roberto Therón,
Francisco J. García-Peñalvo, Martín Martín-González

En este caso de las variantes B y C del cuestionario, el texto de introducción pasaría a ser:

Bienvenido al Cuestionario de Empleabilidad y Empleo Universitarios

Gracias por querer participar. Este cuestionario pretende conocer y analizar la situación de egresados de estudios de máster de universidades españolas como tú. Nuestro objetivo con este análisis es ayudar con los resultados a diseñar políticas educativas y de empleo que mejoren la coordinación entre la enseñanza universitaria y el mercado de trabajo; favoreciendo que los egresados como tú podáis tener más oportunidades de conseguir un empleo o mejorarlo.
Te recordamos una vez más que este cuestionario es anónimo y cumple con lo estipulado en la Ley Orgánica de Protección de Datos de Carácter Personal (LOPD).

También queremos premiar tu dedicación al completarlo, por ello al finalizarlo podrás participar en el sorteo de 10 tabletas Samsung Galaxy Tab 4.

Si tienes cualquier duda o comentario, no dudes en contactarnos en el correo udyc@oeeu.org o en el teléfono 913364185

¡Muchas gracias por colaborar!

Pertinencia							Claridad							Importancia						
1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Observaciones

Autor/es: Juan Cruz-Benito, Roberto Therón,
Francisco J. García-Peñalvo, Martín Martín-González

En las preguntas relacionadas con la comunidad en la que viven, cambiar un selector *drop-down* por un mapa con las comunidades autónomas de España en la que el usuario pueda pinchar en la correspondiente. En este caso se busca que haya elementos visuales diferentes a los habituales y que permitan que el usuario interaccione de formas distintas durante el cuestionario y no sufra tanta fatiga sobre la repetición de acciones.

Como ejemplo, la pregunta:

Pasaría a ser (reconociendo el formulario la comunidad donde pulse el usuario):

Pertinencia							Claridad							Importancia						
1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Observaciones

Inclusión de *feedback* textual relacionado con las respuestas del usuario incluyendo información que pueda ser relevante. Al menos en tres puntos del cuestionario (después de las preguntas de corte demográfico, después de la selección de si se ha estado empleado después del máster, y en la parte final del itinerario escogido). En este cambio se va a introducir una pantalla intermedia entre dos pantallas de preguntas del cuestionario en la que se le dé información al usuario en relación a alguna de sus respuestas frente al resto de las personas que participan en el cuestionario o a estadísticas oficiales de algún tipo.

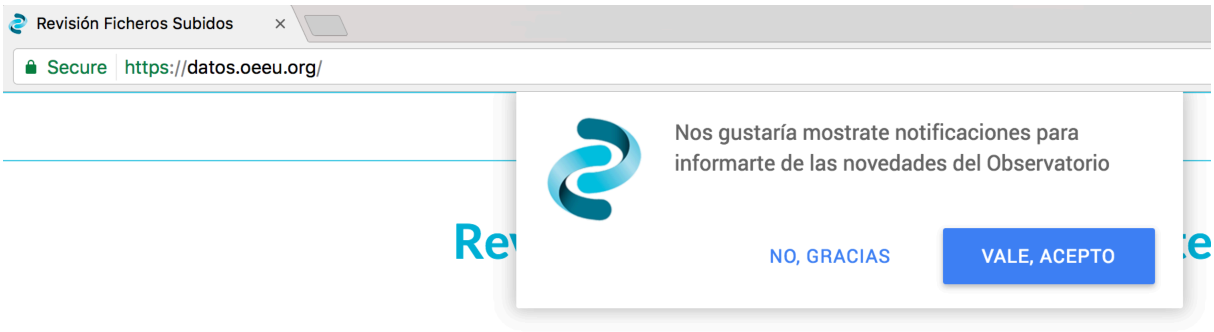
Como ejemplo de este tipo de *feedback*, después de la pantalla del cuestionario donde el usuario responde si ha trabajado alguna vez y cuántos trabajos ha tenido, en el cambio de pantalla de cuestionario (después de pulsar siguiente), aparecería una pantalla solo con una de las siguientes preguntas:

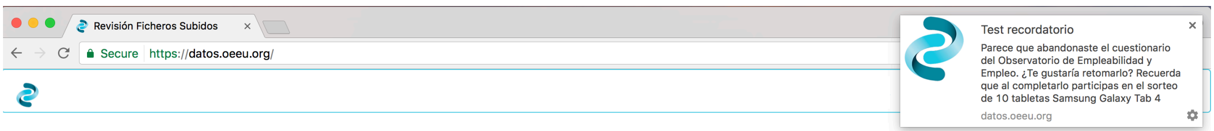
- Si ha respondido que no ha tenido empleo: "¿Sabías que hay un XX% de egresados de máster en tu promoción que tampoco han conseguido un empleo?"
- Si ha respondido que ha tenido un empleo: "¿Sabes que la tasa de empleo de egresados de máster en España es de un XX%?
- Si ha respondido que ha tenido varios empleos: "¿Sabías qué…? Como tú hay un XX% de personas que han respondido a este cuestionario y están en tu misma situación.

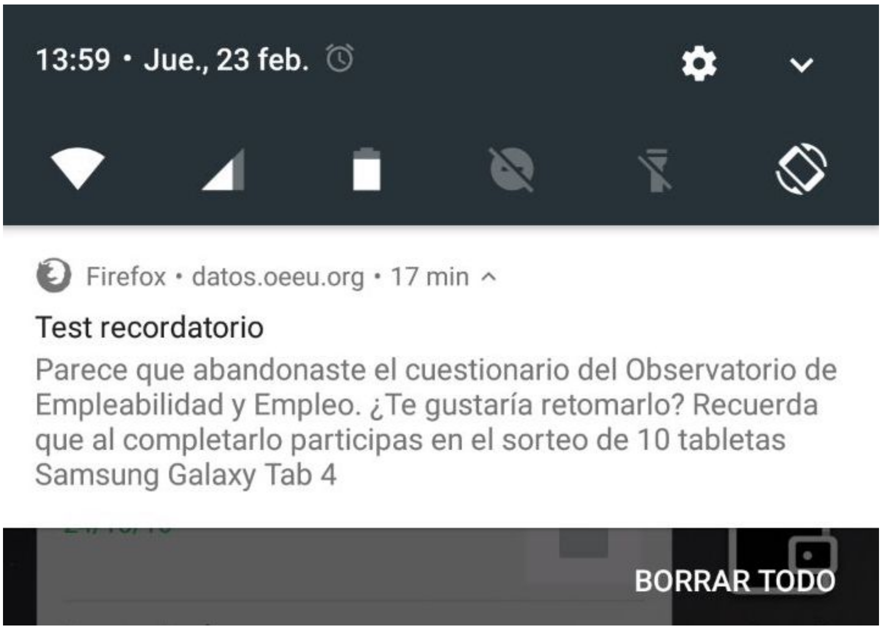
Pertinencia							Claridad							Importancia						
1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Observaciones

Inclusión de notificaciones *push* web que permitan mandar al usuario mensajes de ánimo si abandona el cuestionario. Estas notificaciones solo podrán ser enviadas si el usuario acepta explícitamente. Las notificaciones irán acompañadas del enlace para retomar su cuestionario. Desde un punto de vista técnico, las notificaciones se enviarán a los navegadores Chrome, Firefox y Safari en los Sistemas Operativos Windows, Linux y MacOS de escritorio y a móviles Android con alguno de esos navegadores (cuota total de mercado estimada que se cubre un 61-77%).







Pertinencia							Claridad							Importancia						
1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Observaciones

6.10 Appendix J. Enabling Adaptability in Web Forms Based on User Characteristics Detection Through A/B Testing and Machine Learning

IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received October 23, 2017, accepted November 29, 2017, date of publication December 27, 2017, date of current version February 14, 2018.

Digital Object Identifier 10.1109/ACCESS.2017.2782678

Enabling Adaptability in Web Forms Based on User Characteristics Detection Through A/B Testing and Machine Learning

JUAN CRUZ-BENITO[1], (Member, IEEE), ANDREA VÁZQUEZ-INGELMO[1], JOSÉ CARLOS SÁNCHEZ-PRIETO[2], ROBERTO THERÓN[3], FRANCISCO JOSÉ GARCÍA-PEÑALVO[1], (Member, IEEE), AND MARTÍN MARTÍN-GONZÁLEZ[4]

[1]GRIAL Research Group, Computer Science Department, Research Institute for Educational Sciences, University of Salamanca, 37008 Salamanca, Spain
[2]GRIAL Research Group, Research Institute for Educational Sciences, University of Salamanca, 37008 Salamanca, Spain
[3]GRIAL Research Group, Computer Science Department, and VisUSAL Research Group, University of Salamanca, 37008 Salamanca, Spain
[4]UNESCO Chair in University Management and Policy, Technical University of Madrid, 28003 Madrid, Spain

Corresponding author: Juan Cruz-Benito (juancb@usal.es)

This work was supported in part by the la Caixa Foundation, in part by the Spanish Government Ministry of Economy and Competitiveness throughout the DEFINES Project under Project TIN2016-80172-R, and in part by the University of Salamanca through the Program of Financial Aid for the Predoctoral Contract of José Carlos Sánchez-Prieto (*Programa III: Ayudas para contratos Predoctorales*) cofounded by Banco Santander. The work of J. Cruz-Benito was supported by the European Social Fund and the Consejería de Educación of the Junta de Castilla y León, Spain.

ABSTRACT This paper presents an original study with the aim of improving users' performance in completing large questionnaires through adaptability in web forms. Such adaptability is based on the application of machine-learning procedures and an A/B testing approach. To detect the user preferences, behavior, and the optimal version of the forms for all kinds of users, researchers built predictive models using machine-learning algorithms (trained with data from more than 3000 users who participated previously in the questionnaires), extracting the most relevant factors that describe the models, and clustering the users based on their similar characteristics and these factors. Based on these groups and their performance in the system, the researchers generated heuristic rules between the different versions of the web forms to guide users to the most adequate version (modifying the user interface and user experience) for them. To validate the approach and confirm the improvements, the authors tested these redirection rules on a group of more than 1000 users. The results with this cohort of users were better than those achieved without redirection rules at the initial stage. Besides these promising results, the paper proposes a future study that would enhance the process (or automate it) as well as push its application to other fields.

INDEX TERMS Adaptability, machine learning, user profiles, web forms, clusters, hierarchical clustering, random forest, A/B testing, human-computer interaction, HCI.

I. INTRODUCTION

Understanding what users do within a system is now a fundamental task in the digital world [1]. Most aspects of modern development workflows include users as a centric part of the design and development process of digital products (i.e., user-centered design [2], [3]). Not only knowing what users do (clicks, workflows, interactions, etc.) within a system is valuable for software developers and designers, but these stakeholders must also pay attention to other related-aspects, like user experience, satisfaction, and trust [4]–[8]. Understanding what users do or feel when they use a system is extremely valuable to validate and improve a system. Analyzing users' interactions or their opinion about what they use makes it possible to ascertain the system's strengths or weaknesses regarding users' experience (mostly user interfaces and parts alike) to improve the system based on evidence.

Besides using the analysis of users' interactions and opinions to improve the worst-perceived parts of a system, developers can use these data to build custom or adaptive solutions for different kinds of users [9]–[11]. Using this idea, software engineers could develop versions of the system in which different version are showed to each kind of user. By knowing user profiles and identifying users' behavior and desires, the system could adapt its components to better match users' expectations and likings, and (probably) boost user performance and satisfaction [8], [9], [12].

For a better understanding of the current paper, the context for this experiment is presented. The research has been conducted using a system that belongs to the Spanish Observatory for University Employability and Employment (OEEU in its Spanish acronym) [13]. This observatory gathers data about employment and employability parameters among the Spanish graduates (after they leave the university) to analyze the information they provide and understand what the employment trends and most important employability factors are for this population. To accomplish this mission, the observatory has developed a complex information system [14], [15] that collects and analyzes data to present the insights to the researchers. The system is implemented using the Python language through the Django framework [16] and many other software libraries; it also keeps the information in a MariaDB relational database. To gather data from Spanish universities and students, the OEEU information system has two main tools: one tool is devoted to obtaining students' raw data provided by the university; the other one is a system that generates custom web forms and questionnaires that are to be completed by the graduates after they leave the university. The problem of these web forms is their length, as they typically include between 30 and 70 questions. This second tool for gathering data (the questionnaires) is a centric part in this research.

The goal of this paper is to present a new approach for enabling adaptability in web-based systems using A/B testing methods and user-tracking and machine-learning algorithms that could lead to improving user performance in completing a (large) web form, validating the obtained results through statistical tests. As a secondary goal, the research presented in this paper also aims to produce all machine learning processes in a white-box way, using algorithms and techniques that allow researchers to understand what is happening in every moment. Moreover, to allow readers and other researchers to follow or reproduce the entire process, this paper provide all the code used in the analysis process in Jupyter notebooks available publicly in Github.

The paper has the following structure: section two (Materials and Methods) explains the different algorithms, data, and research framework. Section three (Results) presents the outcomes obtained in the different steps involved in the research: the results regarding the predictive models that provide the most important users' characteristics on completing the web form, those regarding users' profiles found, and those regarding the guidance of users over the different versions of the system to enhance their performance. The fourth section (Discussion) presents different authors' thoughts, proposals, and considerations about this research and its implications, as well as some future works and general conclusions.

II. MATERIALS AND METHODS

This section outlines the materials and methods used for this research. In the case of materials, the data used and the analysis software are described. In the case of the methods, the different steps needed to apply the machine-learning approach to the analysis process as well as the statistics used to prove the validity and significance of the results are presented.

A. MATERIALS

This subsection presents the different materials involved in this research. The materials can be categorized into two main groups: materials related to the experimentation framework and the software tools used to make the proper analysis and support the research process.

The questionnaires and custom web forms included in the OEEU information system gather data from students in two ways: information provided explicitly by the students (the information provided directly) and *paradata* [17]. The paradata from these questionnaires are the auxiliary data that describe the filling process, such as response times, clicks, scrolls, and information about the device used when using the system. All the data used in this research are taken from these two available sources: the raw input tool used by universities and the web forms tool (providing user inputs and their paradata).

Regarding the data used in this research, it is worth noting that to generate the predictive models needed to characterize the main factors that affect users in completing the questionnaires, the authors have chosen only those available before the users began the questionnaire. This is because the research is focused on investigating which factors predetermine participants' success or failure in completing the form, considering all the factors related only to personal context and device and software used to access the web forms. The data about the personal context of the user are provided by the OEEU's system and include information submitted by the university where the user (graduated) studied. All the information that could be used to create the models that predict whether the user will complete the questionnaire (before starting it) is presented in Table 1. Table 1 also explains the data variables used and whether they were valuable for the models. This research has been carried out with a total of 7349 users (all who have some type of experience with the web forms). Of them, the data from 5768 users were considered initially. Finally, data from 3456 users (those resultant after cleaning the data) were used to train and try the machine-learning algorithms (as will be explained in the following section); 1165 users were the cohort introduced in a phase of reinforcement for the questionnaires that validated the rules generated to adapt the web form to users. This number (1165) includes users who did not complete the web form in the first stage as well as users that joined the experiment during the reinforcement and validation phase. Other users (416) only viewed the web forms without starting them. For that reason, were not considered in the experimental report.

TABLE 1. Initial variables gathered from the OEEU information system to build the predictive models of questionnaires' completion.

Name of the variable in the code	Explanation	Type of information that it provides	Was this variable used finally to build the predictive models?
estudiante_id	ID number of student	Personal information	Yes
annoNacimiento	Year of birth	Personal information	No
sexo_id	Gender (male / female)	Personal information	Yes
esEspannol	Is the student Spanish?	Personal information	No
universidad_id	ID of the university where the graduate studied	Personal information	Yes
estudiosPadre_id	Maximum educational level achieved by the graduate's father	Personal information	No
estudiosMadre_id	Maximum educational level achieved by the graduate's mother	Personal information	No
situacionLaboralPadre_id	Current employment status of the graduate's father	Personal information	No
situacionLaboralMadre_id	Current employment status of the graduate's mother	Personal information	No
oficioProfesionPadre_id	Occupation of the graduate's father	Personal information	No
oficioProfesionMadre_id	Occupation of the graduate's mother	Personal information	No
residenciaFamiliar_id	Place of residence of the graduate's family	Personal information	No
residencia_id	Place of residence of the graduate during studies	Personal information	No
idMaster_id	ID number of the master study	Personal information	Yes
especializacionMaster_id	Specialization of the graduate's master	Personal information	No
masterHabilitante	Is an enabling master?	Personal information	No
titularidadMaster_id	Public or not master	Personal information	No
modalidadMaster_id	Modality of the master (online, physical, etc.)	Personal information	No
cursoInicioMaster	Season of the beginning of the master	Personal information	No
cursoFinalizacionMaster	Season of the completion of the master	Personal information	Yes
notaMedia_id	Average grade of the student	Personal information	No
realizacionPracticasMaster	Did the student professionally practice during the master?	Personal information	No
tiempoDuracionPracticasMaster	Time spent by the student in professional practices during the master	Personal information	No
realizacionErasmusMaster	Did the student do an Erasmus stay?	Personal information	No
tiempoDuracionErasmus_id	Time spent by the student in an Erasmus stay	Personal information	No
paisErasmusMaster_id	Country where the student did an Erasmus stay	Personal information	No
viaAccesoMaster_id	Way of accessing the master	Personal information	No
verticalAsignado	Vertical assigned in the A/B testing for the student	Experiment configuration	Yes
cuestionarioFinalizado	Did the student finalize the questionnaire?	Experiment configuration	Yes
numUniversidades	Number of universities involved in the master	Personal information	Yes
numUniversidadesEspannolas	Number of Spanish universities involved in the master	Personal information	Yes
ramaConocimiento_id	Knowledge branch of the master (healthcare, social sciences, engineering, etc.)	Personal information	Yes
realDecreto	Official statement approving of the master studies program	Personal information	Yes
browser_language	Language of the browser used	Device information	Yes
browser_name	Name of the browser used	Device information	Yes
browser_version	Version of the browser used	Device information	Yes
device_pixel_ratio	Device pixel ratio of the browser	Device information	Yes
device_screen_height	Device screen height	Device information	Yes
device_screen_width	Device screen width	Device information	Yes
landscape	Is the device in landscape mode?	Device information	No
os	Operative system of the device	Device information	Yes
os_version	Version of the operative system used	Device information	Yes
portrait	Is the device in portrait mode?	Device information	No
push_notification	Did accept the graduate push notifications for the web form?	Device information	No
push_notification_id	ID number for the push notification subscription	Device information	No
tablet_or_mobile	Is the device tablet or mobile?	Device information	Yes
userAgent	User agent of the device used	Device information	Yes
viewport_height	Height of the window browser	Device information	Yes
viewport_width	Width of the window browser	Device information	Yes

The variables excluded to build the predictive models are those that have more than 10% of their observations with the null value.

The programming language used to conduct all the analyses and calculations was Python. The concrete Python software tools and libraries used to code and execute the different algorithms and statistics were:

- Pandas software library [18]–[20], to manage data structures and support analysis tasks.
- Scikit-learn [21] library, to accomplish the machine learning workflow [22].
- Jupyter notebooks [23]–[25], to develop the Python code used in this research.

All the code developed to analyze user interactions and create machine-learning models, etc. is available at https://github.com/juan-cb/paper-ieeeAccess-2017 [26].

B. METHODS

As found in the bibliography, the concept of A/B testing (also known as bucket testing, controlled experiment, etc.) applied to websites and the Internet could be explained as follows: "show different variations of your website to different people and measure which variation is the most effective at turning them into customers (or people that complete successfully a task in the website, like in this experiment). If each visitor to your website is randomly shown one of these variations and you do this over the same period, then you have created a controlled experiment known as an A/B test" [27]–[29]. In this case, the authors have prepared three different variations, called verticals A, B, and C. In each variation, the authors introduced several changes related to enhancing the users' trustiness, engagement, make the user interface more conversational, etc. All these changes, introduced in the different variations of the web forms (the verticals) used in this research, were proposed by the authors in previous works [30]. These verticals are used as the website variations in which users (students responding to the questionnaires) are meant to test which version is the best regarding the users' performance in the initial stage. To do so, before the experiment, 5768 users were redirected randomly to the different vertical. In the last part of the experiment, the verticals were used to check whether the rules and users' analysis performed during the machine-learning analytics process improve the users' performance in completing the web forms. In this validating phase (which also will be called reinforcement in this paper), 1165 users were redirected to the verticals using the rules generated analyzing the interaction data from the users that acceded randomly to the verticals.

In general, the performed analysis (based on statistics and machine learning) follows common principles in data science regarding data structuration, tidy data approaches, etc. [18], [20], [31]. As stated in the introduction, the machine-learning process has been implemented in a white-box way; thus, the researchers have selected algorithms and methods to make the workflow explainable. This is extremely important, from the authors' point of view, in a research project like this, as it allows humans to provide feedback to the algorithmic process.

Moreover, these main principles, the different details for the analysis pipeline, and methods used in this research are presented.

To find the best models and most accurate parameters, researchers have tried the following approaches:

1. Create predictive models using all the data together. In this approach, researchers tried to use different groups of variables to create the model: all the variables collected from the users, using derived variables (like whether the browser or operative system used to access were modern), etc.
2. Create predictive models using the verticals gap. In this case, researchers generated a predictive model per each vertical of the A/B test. In this case, the most relevant configuration regarding the variables to build up the model in the previous step is included.

Using the most accurate models, the researchers applied all the stages that will be described below (as well as the details for building the predictive models) to generate the different clusters and obtain the rules used to redirect users within the system.

The workflow established (available at https://github.com/juan-cb/paper-ieeeAccess-2017 [26]) is as follows:

1. Retrieve datasets about users from OEEU's information system.
2. Filter the desired fields from the datasets and merge datasets in a single data frame (a data structure like a table).
3. Data cleaning: remove noise data, remove columns (variables) with too many null (*NaN*) values, and remove all users who have only partial information and not all presented in Table 1.
4. Normalize data with the One-hot encoding algorithm for categorical values in columns [22]. To apply the One-hot encoding, researchers used the get_dummies() function from Pandas library, as presented in [26].
5. Considering the data gathered and the kind of variable (labeled) to predict, the algorithm to use must be related to supervised learning. This is because this kind of algorithm makes predictions based on a set of examples (that consist of a labeled training data set and the desired output variable). Moreover, regarding the dichotomous (categorical) character of the variable to predict, the supervised learning algorithm to apply must be based on classification (binary classification, as we have a label of finalization equal to *true* or *false*). According to the authors' previous experience, the possibility of explaining results and the accuracy desired for the classification, a Random Forest classifier algorithm [32] was selected. In this step, the Random Forest algorithm was executed repeatedly, using a custom method based [26] on *GridSearch* functions from Scikit-learn, to determine the best setup for the dataset given (obtaining the most valuable parameters for the execution).
6. With the best configuration found, train the random forest algorithm (with 33.33% of the dataset) and obtain the predictive model.
7. Using the predictive model, obtain the most important features for the predictive model. To obtain these features authors applied *feature_importances_* method from the Random Forest classifier implemented in Scikit-learn library [26].
8. Using the most important features (those that have an importance higher than a custom threshold value of 0.05—the importance score varies between 0 and 1, where 0 is the worst score and 1 the best one), generate clusters applying hierarchical clustering [33]. The reason to use hierarchical clustering is that the algorithm does not require deciding upon the number of clusters to obtain (so, it does not require also to fix previously Euclidean distances and other parameters); it obtains all possible clusters showing the Euclidean distance between them. These clusters represent the groups of users who have participated in the questionnaire according to the most important factors found in the classification.
9. With these clusters, the researchers investigate which clusters exhibit low performance.
10. Using this knowledge about groups of users with low performance and the heuristics observed, software engineers responsible for the OEEU's information system and its web forms could propose changes and

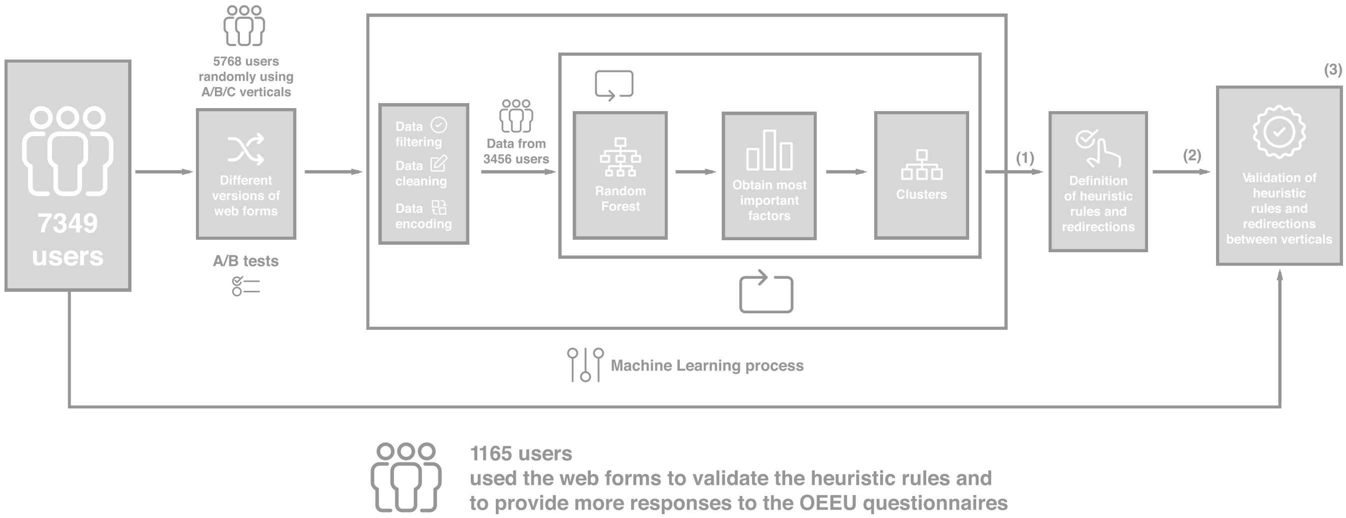


FIGURE 1. Overview of the process followed. Summary of the materials and methods.

fixes (rules, redirections, etc.) in the platform that might help users to improve their performance in the future.

11. Once the data-gathering process is finished, the researchers performed a statistical analysis of the finalization rate of the individuals to determine whether the application of the rules had any impact in the improvement of the finalization of the questionnaires. With this purpose, and considering the characteristics of the variables, the authors applied the Chi-squared test given that it is the most convenient alternative for the analysis of the relationship of two nominal variables.

All these steps and a summary of all methods and materials are presented in the Figure 1.

III. RESULTS

This section presents the main results obtained during the research. The outcomes are divided into three subsections: one related to the results obtained during the machine-learning process (best setup, best ways of building predictive models, the predictive models themselves, the most important variables to finalize or the questionnaire, etc.). The second subsection explains the heuristic rules obtained at the end of the machine-learning workflow inferred from the machine-learning results previously explained. These rules were applied to redirect users within the different verticals of the A/B tests. Finally, the results of the redirections are presented, explaining whether they really affected to the users' finalization of the questionnaire.

A. RESULTS REGARDING MACHINE-LEARNING PROCEDURES: PREDICTIVE MODELS AND CLUSTERING

As previously explained, the researchers performed several attempts to find the most accurate predictive models that better explain whether users will finalize the questionnaire. The first attempt was based on using all the data together focusing in primary variables (excluding those that have too many void values); the second one was based on using all variables and derived variables (constructed from primary ones). The third attempt was based on creating separated predictive models depending on the vertical. In this way, the researchers predicted users' behavior regarding the finalization depending on the vertical / interaction features that they experience. In this last approach, the researchers used the best set of variables found previously to build the model.

The results achieved in this phase would correspond to those expected in the (1) mark in Figure 1.

Regarding the first attempt to build the best predictive model, the researchers used all the variables (excluding the cleaned ones applying the rules defined in the methods sections). As presented in https://github.com/juan-cb/paper-ieeeAccess-2017/blob/master/machinelearning-results.ipynb [26], the predictive model generated had an average precision of 0.79 (Table 2 shows the results and explanations of the results metrics) in predicting whether users will finalize the web form before starting it (in fact, this 0.79 is a fairly good

TABLE 2. Results of the first predictive model built.

	Precision[a]	Recall[b]	F1-score[c]	Support[d]
False	0.84	0.38	0.52	378
True	0.77	0.97	0.86	815
Avg / total	0.79	0.78	0.75	1193

[a] The precision is the ratio *tp / (tp + fp)* where *tp* is the number of true positives and *fp* the number of false positives. The precision is intuitively the classifier's ability of not labeling as positive a sample that is negative. This score reaches its best value at 1 and worst score at 0.
[b] The recall is the ratio *tp / (tp + fn)* where *tp* is the number of true positives and *fn* the number of false negatives. The recall is intuitively the ability of the classifier to find all the positive samples. This score reaches its best value at 1 and its worst score at 0.
[c] The F1 score can be interpreted as a weighted average of the precision and recall, where an F1 score reaches its best value at 1 and its worst score at 0. The relative contribution of precision and recall to the F1 score is equal. This score reaches its best value at 1 and its worst score at 0.
[d] The support is the number of occurrences of each class in each predicted label.

precision score). In the case of this research, the authors use the precision score as the main metric to make decisions, as it is focused on penalizing false positives [34].

The crosstab (that expresses the number of good and bad predictions) for this first predictive model can be found in Table 3.

TABLE 3. Crosstab for the first predictive model built.

	False (predictions)	True (predictions)
False (actual)	142	236
True (actual)	27	788

In this first attempt and its 0.79-precision predictive model, the most important factors in the model were (the importance score varies between 0–1, where 1 is the best score and 0 the worst one):

1. *device_screen_width*: 0.297189
2. *viewport_width*: 0.292615
3. *browser_name_Firefox*: 0.100000
4. *device_pixel_ratio*: 0.098356
5. *viewport_height*: 0.096237

In the second attempt, the researchers used the same variables plus two derived variables composed using the primary ones. The derived variables were *modern_browser* and *modern_os*. Those variables were calculated using the versions of operative systems and browsers used by users. In this case, the researchers calculated the median version of the operative system or browser (the midpoint between the oldest version and newest one present) and classified the browser or operative system as modern or not depending on whether its version is equal or superior to the mid version or is lower. These derived variables were prepared because it was impossible to use the literal version of each browser or operative system in the random forest algorithm due their heterogeneous expressions (each browser or OS has its own version's description and format, etc.). In this second attempt, the precision of the predictive model was higher—specifically, a precision of 0.81 (Table 4). The crosstab for this second model is presented in Table 5.

TABLE 4. Results of the second predictive model built.

	Precision	Recall	F1-score	Support
False	0.91	0.34	0.50	378
True	0.76	0.98	0.86	807
Avg / total	0.79	0.78	0.75	1185

In general, this second model performed better than the previous one (at least it was most precise). In this case, the most important factors that define the model were:

1. *tablet_or_mobile*: 0.179032
2. *device_pixel_ratio*: 0.159406
3. *device_screen_height*: 0.097580
4. *device_screen_width*: 0.095784

TABLE 5. Crosstab for the second predictive model built.

	False (predictions)	True (predictions)
False (actual)	129	249
True (actual)	13	794

5. *viewport_height*: 0.089050
6. *os_Android*: 0.063415

Since the variables used to build the predictive model were different from the previous one, it is normal that the factors that define the model could differ.

In the third approach to generate the best predictive model, the researchers generated a predictive model per each vertical in the A/B test applied to the users. In this case, the researchers included all the variables that produced the best predictive model previously: this is, the variables from the second attempt (including the variables *modern_os* and *modern_browser*). In this case, the researchers have trained three different random forest algorithms, found the best setup for each one depending on the data to analyze, and produced a model for each vertical. The results of these predictive models are presented in Tables 6, 7, and 8, and their precision varied between 0.79 and 0.87. The average precision in the three models was of 0.8233, which is higher than the precision achieved in the previous attempts of generating predictive models. Tables 9, 10, and 11 present the crosstabs for each model; they explain how much effective was the prediction depending on the finalization in the web form.

TABLE 6. Results of the predictive model for the vertical A.

	Precision	Recall	F1-score	Support
False	0.92	0.35	0.51	69
True	0.85	0.99	0.92	263
Avg / total	0.87	0.86	0.83	332

TABLE 7. Results of the predictive model for the vertical B.

	Precision	Recall	F1-score	Support
False	0.92	0.37	0.52	161
True	0.74	0.98	0.85	301
Avg / total	0.81	0.77	0.73	462

Regarding the most important factors per each predictive model generated in the third attempt, the results were the following:

Most influential factors for the predictive model for vertical A:

1. *viewport_width:* 0.267931
2. *tablet_or_mobile:* 0.139438
3. *os_iOS:* 0.132425
4. *device_screen_height:* 0.118814
5. *device_screen_width:* 0.067581
6. *device_pixel_ratio:* 0.066577
7. *os_Android:* 0.054088

TABLE 8. Results of the predictive model for the vertical C.

	Precision	Recall	F1-score	Support
False	0.89	0.37	0.52	132
True	0.74	0.97	0.84	238
Avg / total	0.79	0.76	0.73	370

TABLE 9. Crosstab of the predictive model results for vertical A.

	False (predictions)	True (predictions)
False (actual)	24	45
True (actual)	2	261

TABLE 10. Crosstab of the predictive model results for vertical B.

	False (predictions)	True (predictions)
False (actual)	59	102
True (actual)	5	296

TABLE 11. Crosstab of the predictive model results for vertical C.

	False (predictions)	True (predictions)
False (actual)	49	83
True (actual)	6	232

Most influential factors for the predictive model for vertical B:

1. *viewport_height:* 0.294176
2. *viewport_width:* 0.167701
3. *device_screen_height:* 0.102463
4. *device_pixel_ratio:* 0.085122
5. *os_Android:* 0.076196

Most influential factors for the predictive model for vertical C:

1. *device_screen_width:* 0.193903
2. *viewport_height:* 0.143456
3. *device_screen_height:* 0.108721
4. *tablet_or_mobile:* 0.100000
5. *viewport_width:* 0.093584
6. *device_pixel_ratio:* 0.088479
7. *os_Windows:* 0.055153

Analyzing the results, researchers found that the best way, in this case, to obtain the most-precise predictive models for users' interaction, was obtained by splitting the dataset using the vertical criteria. That is, separating the dataset into three datasets, each one including the data from each user cohort that experienced each one of the A/B tests versions. For that reason, the resultant models were selected to generate the clusters and study them to produce the rules to be used in redirecting users among the different visual representations of the web forms. Using these profiles (clusters) and the rules generated, the researchers found what kind of user (and its technological aspects) fits better (is more inclined to finalize) in each version of the questionnaires, forwarding the users using these criteria to each vertical.

After producing the predictive models, the researchers clustered users depending on their finalization ratio and the most important factors extracted in the predictive models. Explaining all clusters generated after producing each predictive model is out of the scope of this paper (but available at https://github.com/juan-cb/paper-ieeeAccess-2017/blob/master/machinelearning-results.ipynb [26]). Thus, only the clusters obtained after finding the best predictive models will be explained (those generated separately per each vertical). As discussed in the methods section, the clusters were generated using hierarchical clustering techniques because these techniques do not require configuring the target number of clusters. This permits all the relevant clusters (relevance due to the Euclidean distance among them) to be obtained regardless of the number. Figures 2, 3, and 4 present the dendrograms corresponding to each set of clusters.

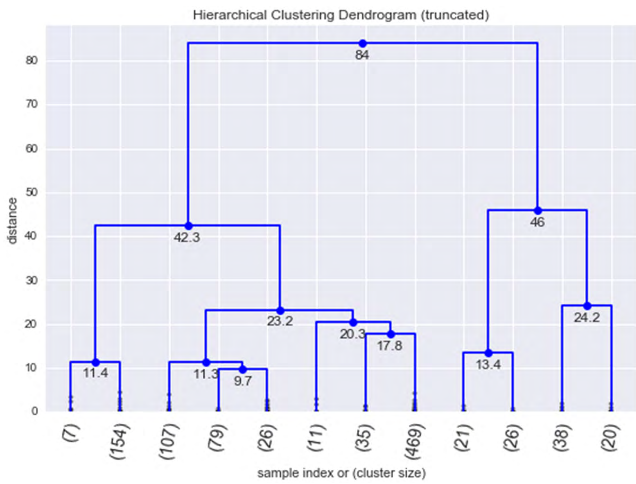


FIGURE 2. Dendrogram representing the clusters found with the predictive model generated using the data from vertical A. Each leaf represents a different cluster obtained (except, in this figure, clusters 8 and 9 that are represented together in the dendrogram due to their closeness in the 9th leaf). The different values that appear near the claves display the Euclidean distance that explains the separation between the different clusters. Finally, the numbers below the leaves (at the bottom of the figure) present the number of users included in the corresponding cluster. Source and full resolution image with all the clusters are available in [26].

After applying the hierarchical clustering algorithm (https://github.com/juan-cb/paper-ieeeAccess-2017/blob/master/machinelearning-results.ipynb [26]) the following numbers of clusters were found: 13 clusters for the vertical A predictive model, 12 clusters for the vertical B model, and 12 clusters for the vertical C.

Analyzing the generated clusters, the researchers found the features that define each cluster and compared them among

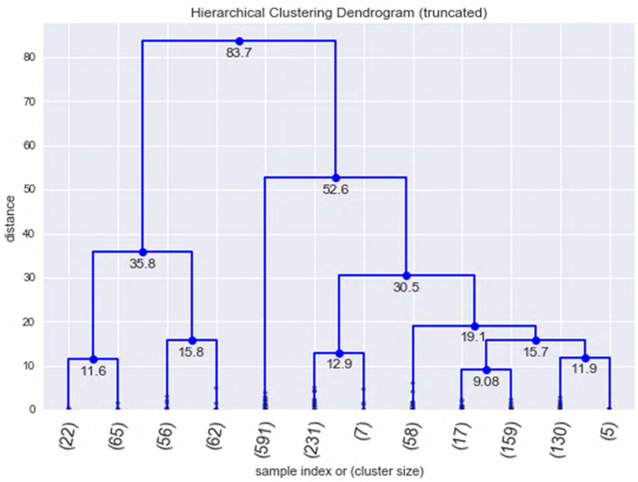


FIGURE 3. Dendrogram representing the clusters found with the predictive model generated using the data from vertical B. The meaning of the different visual elements is the same than those presented in Fig 2. Source [26].

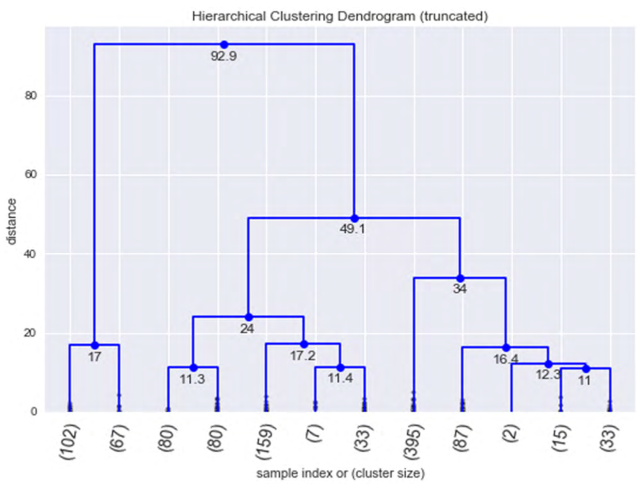


FIGURE 4. Dendrogram representing the clusters found with the predictive model generated using the data from vertical C. The meaning of the different visual elements is the same as those presented in the previous dendrogram figures. Source [26].

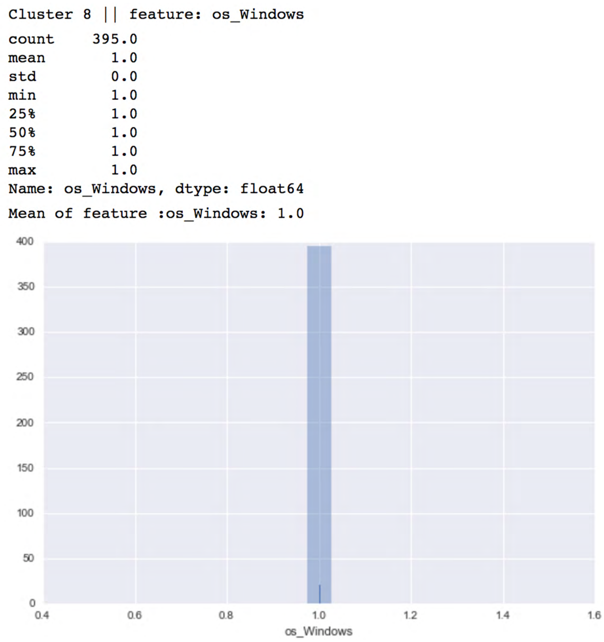


FIGURE 5. Descriptive statistics and distribution of values for cluster 8 (vertical C), regarding the use of the Windows operating system. Source [26].

the different models to define the redirection rules. This analysis of clusters and rule generation will be explained in the following subsection.

B. RESULTS REGARDING CRITERIA FOR REDIRECTING USERS WITHIN A/B TESTING VERTICALS

Once the clusters were identified through the produced predictive models, the researchers started to analyze the features of each cluster to establish the proper redirection rules based on the heuristics observed. In the case of this study, these rules were not generated automatically, although using the code and procedures previously presented, it would be possible. The results achieved at this stage correspond to those expected in mark (2) in Figure 1.

First, the most important values of these features were obtained through descriptive statistics and distribution plots (for every identified cluster), as included in [26]. As an example of the features' identification, Figure 5 shows that in vertical C's 8^{th} cluster, the device's operating system of the clustered users is Windows (the most repeated value is 1, i.e., *True*). With this information (and the rest of information obtained through the same process on the rest of features) the researchers could determine the possible devices used by the students in every cluster. In this case, the authors will refer mainly to these factors as *technical features* or *technical info*, as the factors were all related to the technological aspects of the device and software used by users completing the questionnaires.

The descriptive statistics and distribution plots for every technical feature within each cluster are available at https://github.com/juan-cb/paper-ieeeAccess-2017/blob/master/machinelearning-results.ipynb [26].

Once the values (technical specs mainly) of the devices were obtained, the finalization rates of the questionnaires of all clusters were calculated, identifying the performance achieved by users in each of them. This allowed the identification, for example, of the clusters whose finalization rate were smaller than the finalization rate of the whole questionnaire vertical.

In this way, researchers identified the factors (the most relevant features of each vertical's predictive model) linked to the clusters that performed worse than the rest. This information is summarized in Tables 12, 13, and 16 for verticals A, B, and C, respectively.

These tables (12, 13, and 14) helped the researchers to define the redirection rules. For example, Android devices with a 2-pixel ratio (i.e., Android devices with good screen

TABLE 12. Cluster characteristics identification in vertical A. Clusters that performed below the general completion rate of the vertical are marked in red.

Vertical	Vertical completion rate		Total users		Completed questionnaires			Uncompleted questionnaires		
A	76.23% (average)		993		757			236		
Cluster number	Users count	Completion rate	Viewport width	Tablet or mobile?	iOS?	Screen height (px)	Screen width (px)	Pixel ratio	Android?	Possible device
1	7	71%	2569	False	False	1440	1560	1 or 2	False	Windows computer
2	154	86.36%	1920	False	False	1080	1920	1	False	Windows computer
3	107	83.17%	1440	False	False	900	1440 or 1600	1	False	Windows computer
4	79	79%	1260	False	False	1024	1280	1	False	Windows computer
5	26	80%	1250	False	False	1080	1800	1	False	Windows computer
6	11	81.81%	896 or 1280	False	False	800 or 1024	896 or 1280	1	True	Convertible device
7	35	82.85%	1290	False	False	800 or 900	1280 or 1440	2	False	Retina Mac computer
8	35	80%	1024	False	False	768	1024	1	False	Windows computer
9	434	82.02%	1366	False	False	768	1366	1	False	Windows computer
10	21	9%	366	True	False	640	360	3 or 4	True	Android mobile (very high resolution)
11	26	15%	360	True	False	600 or 700	360	2	True	Android mobile (good resolution)
12	38	2%	500–400	True	True	600	375	2 or 3	False	iPhone
13	20	84.99%	768 or 1024	False	True	1024	768	1 or 2	False	iPad

TABLE 13. Cluster characteristics identification in vertical B. Clusters that performed below the general completion rate of the vertical are marked in red.

Vertical	Vertical completion rate		Total users		Completed questionnaires		Uncompleted questionnaires	
B	66.5% (average)		1403		933		470	
Cluster number	Users count	Completion rate	Viewport height (px)	Viewport width (px)	Screen height (px)	Pixel ratio	Pixel ratio	Possible device
1	22	72.72%	628	414	736	3	False	Large iPhone (iPhone 6 Plus, 6s Plus or iPhone 7 Plus)
2	65	1.5%	450–500 or 550–600	320 or 375	480, 568 or 667	2	False	iPhone
3	56	8.9%	550	360	640	2	True	Android mobile (good resolution)
4	62	11.29%	537	360	640	3-4	True	Android mobile (very high resolution)
5	591	75.8%	649	1366	768	1	False	Windows computer
6	231	73.5%	955	1860	1080	1	False	Windows computer
7	7	71.42%	1290	1960	1440	1	False	Non-retina Mac computer
8	58	77.58%	720	1134	800-900 or 1024	2	False	iPad
9	159	76.72%	620	1260	1024, 1080 or 1200	1	False	Windows computer
10	17	76.47%	780	1440 or 1600	900	1	False	Windows computer
11	130	74.61%	894	1260	1024	1	False	Windows computer
12	5	80%	936 or 1144	768-800	1024 or 1080	1	True	Android tablet

resolution), despite their low rate performance, obtain better finalization ratios in vertical A (finalization rate of 15%) than in verticals B and C (finalization rates of 8.9% and 10.7%, respectively), leading to the conclusion that the users with devices that meet these characteristics should be redirected to vertical A.

Repeating this methodology for every device identified, the following rules were obtained:

TABLE 14. Cluster characteristics identification in vertical C. Clusters that performed below the general completion rate of the vertical are marked in red.

Vertical	Vertical completion rate		Total users		Completed questionnaires			Uncompleted questionnaires		
C	65% (average)		1060		689			371		
Cluster number	**Users count**	**Completion rate**	**Screen width (px)**	**Viewport height (px)**	**Screen height (px)**	**Tablet or mobile?**	**Viewport width (px)**	**Pixel ratio**	**Windows?**	**Possible device**
1	102	10.7%	360	550	640	True	360	2	False	Android mobile (good resolution)
2	67	20.84%	360	570	640	True	370	3	False	Android mobile (very high resolution)
3	80	71.25%	1280	895	1024	False	1280	1	True	Windows computer
4	80	77.5%	1440, 1600 or 1920	760	900	False	1440 or 1600	1	True	Windows computer
5	159	72.95%	1920	950	1080	False	1920	1	True	Windows computer
6	7	71.42%	2560	1240	1440	False	1892	1	False	iMac
7	33	72.72%	1920	928	1080	False	1700	1	False	Non-retina Mac computer
8	395	76.96%	1366	645	768	False	1366	1	True	Windows computer
9	87	71.26%	1350	670	800–900	False	1280-1300	1	False	Non-retina Mac computer
10	2	50%	1080	500	1848	True	360	3	False	Android tablet
11	15	66.66%	768	950	1024	False	768	1 or 2	False	Mac computer
12	33	69.69%	1148	1280	800 or 1024	False	1280	2	False	Retina Mac computer

- Redirection to vertical A:
  - Android devices with a 2-pixel ratio.
  - Computers with an operating system different from Android, iOS and Mac OS.
  - Mac OS computers.
  - iPad devices.
  - Convertible devices (those that could be used as tablet or as laptop depending on whether a keyboard or mouse is attached to them).
- Redirection to vertical B:
  - Android devices with a 3- or 4-pixel ratio.
  - Large iPhone devices (iPhone 6 Plus, 6s Plus, or 7 Plus).
  - Android tablets.

If the devices of the users who participate in the reinforcement (validate) phase did not meet any of these characteristics, the redirection was randomly made between verticals A and B (maintaining a 50% distribution).

No users were redirected to vertical C due to the low finalization rates of the clusters in this questionnaire variant. There was only one rule that did not follow this assumption: the case of an Android device with a very high resolution (a pixel ratio of 3 or 4). Despite this case, the researchers decided to close this vertical C, as all the mobile or tablet devices with a very high resolution (like iPhone 6 Plus, 6s Plus, 7 Plus, or Android tablets) work better in vertical B.

The final established heuristic rules were the following (presented as a kind of pseudocode):

1. If the operating system is Android and the device's pixel ratio is 2, the user is redirected to vertical A.
2. If the operating system is Android and the device's pixel ratio is 3 or 4, the user is redirected to vertical B.
3. If the operating system of the device is iOS and its pixel ratio is 3 (iPhone 6 Plus, 6s Plus, or 7 Plus), then the user is redirected to vertical B.
4. If the operating system is neither Android nor Mac OS, iOS, the user is redirected to vertical A.
5. If the operating system of the device is Mac OS, the user is redirected to vertical A.
6. If the operating system is Android and the device's screen height is greater than 1000px, the user is redirected to vertical B.
7. If the operating system is iOS, the device's screen width is 1024px, the device's screen height is 768px, and the device's pixel ratio is 1 or 2 (iPad), the user is redirected to vertical A.
8. If the device's operating system is Android and the device type is neither a mobile nor a tablet (convertible device), the user is redirected to vertical A.
9. If a device does not fit any of the previous conditions, the user is randomly redirected to vertical A or B (with equal probability of being redirected to any of them).

These rules were implemented in the OEEU's ecosystem to apply them whenever a new user enters or resumes the questionnaire.

C. RESULTS REGARDING ADAPTABILITY AND USERS' REDIRECTION WITHIN A/B TEST VERTICALS

After the experiment took place (analyzing the interaction and performance of users who used the system previously), all the users who entered or returned to the questionnaire (and therefore, the target users of the experiment) were sought to obtain the results regarding the application of redirection criteria within the questionnaire verticals. The calculation and validation presented in this phase correspond to the (3) mark in Figure 1.

Before this phase (called reinforcement because the participants are users who access the web forms in a reinforcement made by the OEEU to obtain more responses to the questionnaires) and the application of the redirection rules based on heuristics, 5768 users had started the questionnaire; 4410 of them finished it, leaving a total of 1358 uncompleted questionnaires (and reaching a completion rate of 76.46%). All the data related to this subsection are available at https://github.com/juan-cb/paper-ieeeAccess-2017/blob/master/reinforcement-results.ipynb [26]

In these previous results, the users who *entered* the questionnaire (i.e., reached the welcome page but never started it) were not taken into account. If these users were considered, the results would be as follows:

- Number of students who have *entered* the questionnaire: 6360.
- Number of students who have *not finished* the questionnaire: 1950.
- Number of students who have *finished* the questionnaire: 4410.
- Completion rate *before* reinforcement: 69.34%.

By the time the questionnaires were closed, the final results were the following: 6738 started questionnaires, of which 5214 were completed and 1524 uncompleted. Consequently, the study achieved a questionnaire completion rate of 77.38%, improving the previous rate.

Again, these are the results for the started questionnaires; considering all the users (including the ones who reached the welcome page), the study yields the following results:

- Number of students who have *entered* the questionnaire: 7349.
- Number of students who have *not finished* the questionnaire: 2135.
- Number of students who have *finished* the questionnaire: 5214.
- Completion rate *after* reinforcement: 70.95%.

The total number of target users who entered the questionnaire after the incorporation of the system redirection support was 1165. These 1165 users were classified into three groups:

- Users who *entered* the questionnaire *after* reinforcement (considered as "new users"). There were 1003 new users, becoming the larger group of users who have taken part in the experiment.
- Users who resumed the questionnaire *after* reinforcement and were redirected to a different vertical; 110 users satisfied this criterion.
- Users who resumed the questionnaire *after* reinforcement but were *not* redirected to a different vertical. There were 52 users of this type.

These general results are summarized in Table 15.

TABLE 15. General results in the reinforcement phase.

User type	Total	Results	Completion rate
New users	1003	718 finished questionnaires 285 *not* finished questionnaires	71.59%
Redirected users	110	61 finished questionnaires 49 *not* finished questionnaires	55.45%
Not redirected users	52	25 finished questionnaires 27 *not* finished questionnaires	48.08%

As can be seen in Table 15, the new users' sample reached a completion rate of 71.59%.

This sample includes users who (at least) reached the welcome page of the questionnaire after reinforcement. An improvement in the completion results could be seen when comparing this completion rate (71.59%) with the completion rate before the reinforcement (that includes all the users who entered the questionnaire, 69.34%). Furthermore, it is necessary to consider that these new users are more reluctant in completing the questionnaire, as they have been invited to participate at least twice previously (and they had ignored the invitations), so these results are even more valuable.

Once the participant finalization rates were calculated, the researchers proceeded with the analysis of the impact of the rules formulated to improve the finalization rate, taking as a reference the groups of users who accessed the questionnaire presentation page both before and after the reinforcement phase.

These users were grouped into categories according to the way in which they were assigned to their vertical. To generate these categories, the researchers applied the assignment rules to the group of users who participated prior to the reinforcement and compared the results (ideal vertical assignment) with the vertical to which these individuals were actually sent (actual vertical assigned). Thus, the following three groups of individuals were obtained:

- **Pre-reinforcement users randomly assigned to the wrong vertical (G1, n = 3833):** Composed of users who

TABLE 16. Correlation between the vertical assignment and the finalization rate.

	Finalization rate	Chi-squared	Significance
G1-G2	67.9-74.9	25.927	0.000
G2-G3	74.9-71.6	3.442	0.064
G1-G3	67.9-71.6	5.130	0.024

accessed the questionnaire before the reinforcement and were assigned to a vertical to which they would not have been assigned had the redirection rules been applied.

- **Pre-reinforcement users randomly assigned to the right vertical (G2, n = 1542):** Comprised of users who accessed the questionnaire before the reinforcement and who, despite having been randomly directed, were assigned to the vertical to which they would have belonged to, had the redirection rules been applied.
- **Post-reinforcement users (G3, n=1003):** Users who accessed the questionnaire for the first time after the reinforcement, thus being consequently assigned to the right vertical.

In the case of rule 9, researchers classified all individuals who were randomly directed to vertical C as members of group 1; individuals who were directed to verticals A or B were classified as missing values, as the distribution of those verticals was defined differently from the one defined for the reinforcement phase.

Once the users were classified, the researchers calculated the finalization rate of each group, using the Chi-square statistic to study whether the vertical assignation method influenced the finalization rate. The Chi-square test is the most reliable in this scenario, given that there are two categorical variables (questionnaire finalization and success in the assignment). This statistical test was applied to the three possible combinations of pairs (Table 16).

First, as we can observe in the table, the results of the Chi-square test reflect a significant correlation between the vertical assignation method and the finalization of the questionnaire in pair G1–G2 for a significance level (s.l.) of 0.001. This result is consistent with the methodology employed, given that the clustering process and the later rules of assignment were carried out using the pre-reinforcement users.

Second, for the pair G2–G3, the results indicate no correlation between the assignment method and the finalization rate (s.l. 0.05) which, again, confirms the adequacy of the established rules, as individuals in group 3 were grouped with the same criterion that those in group 2, although the assignment was done in an intentional way rather than randomly.

Finally, it is noticeable that there is also a correlation (s.l. 0.05) between the assignment method and the finalization rate in the case of the pair G1–G3, which confirms that the application of the established rules significantly contributes to the finalization of the questionnaire by the participants.

TABLE 17. Correlation between the application rule and the finalization rate.

	Finalization rate		N		Chi-squared	Significance
	G1	G3	G1	G3		
Rule 1	67.11	72.64	374	106	1.167	0.280
Rule 2	70.99	72.22	362	126	0.069	0.793
Rule 3	62.75	56.52	51	23	0.258	0.612
Rule 4	68.03	76.25	2196	421	11.215	0.001
Rule 5	71.69	73.47	325	49	0.067	0.796
Rule 6	*	*	*	*	*	*
Rule 7	67.12	74.07	73	27	0.445	0.505
Rule 8	72.00	80.00	25	10	0.643**	0.488**
Rule 9	62.53	63.07	427	241	0.019	0.890

*No individuals in group 2. **Fisher's exact test (odds ratio and p-value)

As a final data analysis step, the researchers carried out an in-depth study of the behavior of each of the proposed rules, aiming to delve into the individual effect of each of them on the finalization of the questionnaire.

To this end, a process like the previous analysis was used with each one of the rules, the difference being that only the pair G1–G3 was used (Table 17).

As illustrated in Table 16, although there are differences in all finalization rates, they are significant (s.l. 0.01) only in the case of rule 4. For said rule, the rate of finalization in group 3 is approximately 8% greater than the rate in group 1, which suggests that directing the individuals who access the questionnaire from a non-Mac PC improves their chances of completing the questionnaire.

IV. DISCUSSION

This section presents the discussion of all the issues found in the research, discussing the foundations and effects of some decisions made by the authors. It also includes several future lines of work, suggests a set of recommendations, and closes with a general conclusion.

A. GENERAL DISCUSSION

Regarding the research carried out by the authors, there are several issues to comment on in this paper. To facilitate the comprehension, these issues will be discussed following the same structure of the paper (first, issues related to the methodology; second, those related to the results, and so forth).

First, the authors pose a question: Is it advisable to apply this kind of machine-learning method to this kind of problem? In this case, the researchers were inspired by other authors who have applied these types of processes to a wide range of problems. As an example, this kind of machine-learning algorithmic approach has being used in other fields, such as education [35], with promising results. Beyond the benefits that machine-learning approaches bring to many problems, by also including white-box procedures, the researchers ensure explainable and reproducible results that could be improved or discussed by the scientific community. All these considerations and precedents encouraged the authors to employ this kind of approach to address the problem of improving users' performance within a complex system like that presented. According to the results, the question can be answered positively, as the findings have been valuable and prove the validity of the approach.

Following the discussion, the authors would like to comment that the A/B testing approach used for this research is not a *pure* application of such methodology. While A/B tests are commonly based on singular changes between the different experimentation groups (or verticals), in the presented approach the authors grouped different changes into the same verticals. In this case, this variation of A/B tests does not influence this experimentation, as the researchers attempt to maximize user performance in the questionnaire finalization without a special focus on small changes, but using important differences between the different verticals. Despite that, it is worth noting that this kind of application of A/B tests for the experiment has been previously validated by experts [29].

Regarding the generated predictive models, the cut-off value for their relevant factors to later include in the clusters, the authors stated 0.05 as the minimum value to consider since this is the most common value in classical literature to ensure reliable results. Also in this case, the authors use this cut-off value to generate the clusters using only the most important factors (those that have a specific weight of more than 0.05 in the predictive model), thus excluding less important ones that could introduce noise when building the groups.

Concerning the most important factors that characterize the predictive models and explain the users' profile and preferences while completing the questionnaire, it should be remarked that technical aspects were more important than personal ones. At the beginning of the research and for the predictive models' generation, researchers included personal aspects, such as gender, age, and issues related to education, as part of the dataset. According to the results, such aspects do not have special relevance while modeling the users' behavior in completing the web form. Instead, the present findings indicated that the most important factors for the users were the size of the device screen and the browser window. Moreover, other aspects, like the screen resolution, concrete browser, or operative system, were important, but with a lesser effect. Nevertheless, these are the most important factors for the population of this study and cannot be considered general and valid for other populations. To apply the approach presented in this research in other experiments, the predictive models should be generated again.

Regarding the generation of rules based on heuristics, and as a future study, the researchers would like to automate this process. This will help to reproduce the same process with the same experimental conditions and remove any kind of bias introduced by researchers or administrators. This will be explained in depth in the following subsection.

Related to the reinforcement phase and other conditions of the experiment, with the aim of enhancing users' participation in the questionnaires, the OEEU offered participation in a raffle (the prize would be seven smartwatches) to all graduates completing the web form as a reward. This incentive was used also to promote the reinforcement process where the redirection rules were applied.

Regarding the effectiveness of the use of rules based on cluster analysis during the reinforcement period, cluster analysis was found to be a very useful tool to guide the redirection of users to the version of the questionnaire best suited to the features of the technology with which they completed it.

First, the results of this study confirm that the rules established improved the answer rate by comparing the performance of users who participated after the reinforcement with those who participated before the last reinforcement and were directed to the wrong questionnaire. Additionally, the authors could observe that there are no significant differences between groups G2 and G3, which leads to the understanding that the application of the rules during the reinforcement has maintained the good results regarding to the finalization among the users who would have been randomly assigned to the right vertical.

Second, if the researchers delve into the analysis of the individual behavior of each rule, the results suggest that the improvement in the finalization rate is due to rule four, which redirects users who access the form from non-Mac computers to vertical A, given that the rest of rules have not yielded significant correlations.

Regarding this point, it must be remarked that the users who participated in a reinforcement phase were commonly more reluctant to complete the questionnaire, as they left it in previous stages or were not initially attracted to fulfill it. This also could render even more valuable the results obtained in this research concerning the improvement of users' performance. However, for future studies it would be interesting to apply a research design that includes an experimental and a control group from the beginning to be able to assess the effect of the rules under the same conditions.

Another interesting future line of research would be an analysis of the threshold cut-off to perform the factor selection, given that a higher minimum value may simplify the number of rules and make more efficient the redirection process. As a first step, the authors intend to analyze rule four to gain a better understanding of the predictive importance of the elements behind its formulation.

Finally, the authors believe that the approach and procedures presented in this research are transferable to other application fields. The process presented in this paper follows some *traditional* approaches and methods within the machine-learning research field, and the prediction challenge is present in many other problems beyond web form completion. The proposed methodology may also help to transfer this experience to other problems with the additional value of providing a white-box approach for the algorithms used. In the future, the authors would like to attempt to apply such methodology to predict the employability of Spanish graduates. This will also validate the genericity of the methodology, which will only require some minor changes depending on the dataset.

B. HOW TO APPLY THIS RESEARCH IN PRODUCTION IN THE REAL WORLD

One of the main concerns related to this research could be stated as follows: Is it possible to use this contribution in a real industry setting? Is it possible to integrate this kind of approach in production systems and enable an automated process? From the point of view of the researchers, the answer is yes to both questions. There are many examples in the industry on how data sciences processes can be transformed from Jupyter notebooks to enterprise-ready systems put in production. In this case, the researchers outline the approach proposed by the Airbnb engineering team on how their ML Automator [36] tool helped in translating a Jupyter notebook into an Airflow machine learning pipeline [37] and use this kind of analytics process in production systems. This automating effort must include—apart from the machine-learning algorithms and process—rule generation or the identification of the proper Euclidean distance to separate the clusters generated. To automate the rule generation, probably researchers would have to employ artificial intelligence techniques such as neuronal networks, that could learn to generate these rules as done by humans in this paper.

C. GENERAL CONCLUSION

This paper presents a novel study in the field of Human-Computer Interaction. The main results achieved have been quite promising and encourage authors to continue the labor of improving users' performance in completing large web forms. Adaptability can be achieved by detecting users' behaviors, preferences, and profiles using machine-learning techniques and offering the best user interface and user experience to each kind of user detected. Based on the results, the authors also propose several future works that could push this research to be adopted in the industry and other application fields.

ACKNOWLEDGMENTS

The authors thank Felicidad García-Sánchez for designing Figure 1.

REFERENCES

[1] R. Colomo-Palacios, F. J. García-Peñalvo, V. Stantchev, and S. Misra, "Towards a social and context-aware mobile recommendation system for tourism," *Pervasive Mobile Comput.*, vol. 38, pp. 505–515, Jul. 2017.
[2] D. A. Norman and S. W. Draper, Eds., *User Centered System Design: New Perspectives on Human-Computer Interaction*. Hillsdale, NJ, USA: L. Erlbaum Assoc. Inc., 1986.
[3] D. A. Norman, *The Design of Everyday Things: Revised and Expanded Edition*. New York, NY, USA: Basic Books, 2013.
[4] J. A. Bargas-Avila, O. Brenzikofer, S. P. Roth, A. N. Tuch, S. Orsini, and K. Opwis, "Simple but crucial user interfaces in the World Wide Web: Introducing 20 guidelines for usable Web form design," in *User Interfaces*, R. Matrai, Ed. Rijeka, Croatia: InTech, 2010. [Online]. Available: https://www.intechopen.com/books/user-interfaces/simple-but-crucial-user-interfaces-in-the-world-wide-web-introducing-20-guidelines-for-usable-web-fo, doi: 10.5772/9500.
[5] B. Shneiderman and C. Plaisant, *Designing the User Interface*, 4th ed. Reading, MA, USA: Addison-Wesley, 2005.
[6] M. Seckler, S. Heinz, S. Forde, A. N. Tuch, and K. Opwis, "Trust and distrust on the Web: User experiences and website characteristics," *Comput. Human Behavior*, vol. 45, pp. 39–50, Apr. 2015.
[7] B. Shneiderman, "Designing trust into online experiences," *Commun. ACM*, vol. 43, no. 12, pp. 57–59, Dec. 2000.
[8] S. P. Anderson, *Seductive Interaction Design: Creating Playful, Fun, and Effective User Experiences, Portable Document*. London, U.K.: Pearson Education, 2011.
[9] S. Krug, *Don't Make Me Think: A Common Sense Approach to Web Usability*. Bengaluru, India: Pearson Education, 2000.
[10] D. Malandrino, F. Mazzoni, D. Riboni, C. Bettini, M. Colajanni, and V. Scarano, "MIMOSA: context-aware adaptation for ubiquitous Web access," *Pers. Ubiquitous Comput.*, vol. 14, no. 4, pp. 301–320, May 2010.
[11] G. Pruvost, T. Heinroth, Y. Bellik, and W. Minker, "User interaction adaptation within ambient environments," in *Next Generation Intelligent Environments*. New York, NY, USA: Springer, 2016, pp. 221–263.
[12] C. Flavián, R. Gurrea, and C. Orús, "The effect of product presentation mode on the perceived content and continent quality of Web sites," *Online Inf. Rev.*, vol. 33, no. 6, pp. 1103–1128, 2009.
[13] F. Michavila, J. M. Martínez, M. Martín-González, F. J. García-Peñalvo, and J. Cruz-Benito, "Barómetro de Empleabilidad y Empleo de los Universitarios en España, 2015 (Primer informe de resultados)," Universidad Politécnica Madrid, Madrid, Spain, 2016.
[14] F. Michavila, M. Martín-González, J. M. Martínez, F. J. García-Peñalvo, and J. Cruz-Benito, "Analyzing the employability and employment factors of graduate students in Spain: The OEEU Information System," presented at the 3rd Int. Conf. Technol. Ecosyst. Enhancing Multiculturality (TEEM), Porto, Portugal, 2015.
[15] A. Vázquez-Ingelmo, J. Cruz-Benito, and F. J. García-Peñalvo, "Scaffolding the OEEU's data-driven ecosystem to analyze the employability of Spanish graduates," in *Global Implications of Emerging Technology Trends*, F. J. García-Peñalvo, Ed. Hershey, PA, USA: IGI Global, 2018.
[16] Django Software Foundation. (Mar. 15, 2015). *Django Web Framework*. [Online]. Available: https://www.djangoproject.com/
[17] S. Stieger and U.-D. Reips, "What are participants doing while filling in an online questionnaire: A paradata collection tool and an empirical study," *Comput. Human Behavior*, vol. 26, no. 6, pp. 1488–1495, Nov. 2010.
[18] W. McKinney, *Python for Data Analysis: Data Wrangling With Pandas, NumPy, and IPython*. Sebastopol, CA, USA: O'Reilly Media, Inc., 2012.
[19] W. McKinney. (2017). *Pandas, Python Data Analysis Library*. Accessed: 2017. [Online]. Available: http://pandas.pydata.org/
[20] W. McKinney, "Data structures for statistical computing in python," in *Proc. 9th Python Sci. Conf.*, vol. 445. Austin, TX, USA, 2010, pp. 51–56.
[21] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Oct. 2011.
[22] S. Raschka, *Python Machine Learning*. Birmingham, U.K.: Packt Publishing Ltd., 2015.
[23] T. Kluyver *et al.*, "Jupyter notebooks—A publishing format for reproducible computational workflows," in *Proc. ELPUB*, 2016, pp. 87–90.
[24] M. Ragan-Kelley *et al.*, "The Jupyter/IPython architecture: A unified view of computational research, from interactive exploration to communication and publication," in *Proc. AGU Fall Meeting Abstracts*, 2014.
[25] F. Perez and B. E. Granger, "Project Jupyter: Computational narratives as the engine of collaborative data science," Tech. Rep., 2015. [Online]. Available: https://blog.jupyter.org/project-jupyter-computational-narratives-as-the-engine-of-collaborative-data-science-2b5fb94c3c58
[26] J. Cruz-Benito, A. Vázquez-Ingelmo, and J. C. Sánchez-Prieto. (2017). *Enabling Adaptability in Web Forms Based on User Characteristics Detection Through A/B Testing and Machine Learning*, Code Repository That Supports the Research Presented in the Paper. [Online]. Available: https://github.com/juan-cb/paper-ieeeAccess-2017
[27] Y. Xu, N. Chen, A. Fernandez, O. Sinno, and A. Bhasin, "From infrastructure to culture: A/B testing challenges in large scale social networks," presented at the 21th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, Sydney, NSW, Australia, 2015.
[28] E. Dixon, E. Enos, and S. Brodmerkle, "A/b testing of a webpage," U.S. Patent 7 975 000, Jul. 5, 2011.
[29] D. Siroker and P. Koomen, *A/B Testing: The Most Powerful Way to Turn Clicks Into Customers*. Hoboken, NJ, USA: Wiley, 2013.

[30] J. Cruz-Benito *et al.*, "Improving success/completion ratio in large surveys: A proposal based on usability and engagement," in *Proc. 4th Int. Conf Learn. Collaboration Technol. Technol. Edu. (LCT)*, 2017, pp. 352–370.

[31] H. Wickham, "Tidy data," *J. Statist. Softw.*, vol. 59, no. 10, pp. 1–23, 2014.

[32] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001.

[33] S. C. Johnson, "Hierarchical clustering schemes," *Psychometrika*, vol. 32, no. 3, pp. 241–254, Sep. 1967.

[34] Scikit-Learn. (Sep. 4, 2017). *API Reference—Scikit-Learn Documentation: Metrics*. [Online]. Available: http://scikit-learn.org/stable/modules/classes.html

[35] A. Zollanvari, R. C. Kizilirmak, Y. H. Kho, and D. Hernández-Torrano, "Predicting students' GPA and developing intervention strategies based on self-regulatory learning behaviors," *IEEE Access*, vol. 5, pp. 23792–23802, 2017.

[36] R. Chang. (Sep. 4, 2017). *Using Machine Learning to Predict Value of Homes On Airbnb*. [Online]. Available: https://medium.com/airbnb-engineering/using-machine-learning-to-predict-value-of-homes-on-airbnb-9272d3d4739d

[37] M. Beauchemin. (Sep. 4, 2017). *Airflow: A Workflow Management Platform*. [Online]. Available: https://medium.com/airbnb-engineering/airflow-a-workflow-management-platform-46318b977fd8

JUAN CRUZ-BENITO received the M.Sc. degree in intelligent systems from the University of Salamanca, Spain, in 2013, where he is currently pursuing the Ph.D. degree in computer sciences. He is one of the youngest members of the Research Group Interaction and eLearning, where he specializes in software solutions based on technology ecosystems and open source software. He is involved in human–computer interaction, educational virtual worlds and technologies for educational purposes, and disciplines that he has developed in many innovation and research projects. He has participated in many European and national research and development projects, such as TRAILER, VALS, USALSIM Virtual Campus, and the Spanish Observatory for University Employability and Employment, as a Software Engineer, Researcher, and Developer.

ANDREA VÁZQUEZ-INGELMO was born in Salamanca, Castilla y León, Spain, in 1994. She received the bachelor's degree in computer engineering from the University of Salamanca, Salamanca, in 2016, where she is currently pursuing the master's degree in computer engineering. She is a member of the Research Group of Interaction and eLearning. Since 2016, she has been a part of the National Project Spanish Observatory for University Employability and Employment as a Developer and a Researcher.

JOSÉ CARLOS SÁNCHEZ-PRIETO received the bachelor's degree in pedagogy and the master's degree in ICT applied in education from the University of Salamanca, Spain. He is currently pursuing the Ph.D. degree with the Faculty of Education, Said University, within the Program on Education in the Knowledge Society. His area of research is the assessment of attitudes among in-service and pre-service teachers.

ROBERTO THERÓN received the Diploma degree in computer science from the University of Salamanca, the B.A. degree from the University of A Coruña, the bachelor's degree in communication studies and the bachelor's degree in humanities from the University of Salamanca, and the Ph.D. degree from the Research Group Robotics, University of Salamanca. His Ph.D. thesis was on parallel calculation configuration space for redundant robots. He is currently the Manager of the VisUsal Group (within the Recognized Research Group GRIAL), University of Salamanca, which focuses on the combination of approaches from computer science, statistics, graphic design, and information visualization to obtain an adequate understanding of complex data sets. He has authored of over 100 articles in international journals and conferences. In recent years, he has been involved in developing advanced visualization tools for multidimensional data, such as genetics or paleo-climate data. In the field of visual analytics, he develops productive collaborations with groups and institutions internationally recognized as the Laboratory of Climate Sciences and the Environment, France, or the Austrian Academy of Sciences, Austria. He received the Extraordinary Doctoral Award for his Ph.D. thesis.

FRANCISCO JOSÉ GARCÍA-PEÑALVO received the degrees in computing science from the University of Salamanca and the University of Valladolid, and the Ph.D. degree from the University of Salamanca. He is currently the Head of the Research Group Interaction and eLearning. He was the Vice Chancellor of Innovation with the University of Salamanca from 2007 to 2009. He has led and participated in over 50 research and innovation projects. He has authored over 300 articles in international journals and conferences. His main research interests focus on eLearning, computers and education, adaptive systems, Web engineering, semantic Web, and software reuse. He is a member of the Program Committee of several international conferences and a reviewer of several international journals. He was a Guest Editor of several special issues of international journals, such as *Online Information Review*, *Computers in Human Behavior*, and *Interactive Learning Environments*. He is currently the Editor-in-Chief of the *International Journal of Information Technology Research* and the *Journal of Education in the Knowledge Society*. He is also the Coordinator of the multidisciplinary Ph.D. Program on education in the Knowledge Society.

MARTÍN MARTÍN-GONZÁLEZ received the master's degree in economic development and public policy and the Ph.D. degree in economics from the Autonomous University of Madrid (UAM), and the degree in economics from the University of La Laguna. He was a Researcher with the Faculty of Economics, UAM. He is currently a Researcher with the UNESCO Chair in University Management and Policy, Technical University of Madrid. His areas of research are the economics of education, the economics of higher education, employability, the evaluation of educational policies, public economics, applied economics, economic development, higher education, and vocational training.

6.11 Appendix K. How different versions of layout and complexity of web forms affect users after they start it? A pilot experience

J. Cruz-Benito, J. C. Sánchez-Prieto, A. Vázquez-Ingelmo, R. Therón, F. J. García-Peñalvo and M. Martín-González, "How different versions of layout and complexity of web forms affect users after they start it? A pilot experience," in World Conference on Information Systems and Technologies, 2018, pp. 971-979: Springer

How different versions of layout and complexity of web forms affect users after they start it? A pilot experience

Juan Cruz-Benito[1,3,4], José Carlos Sánchez-Prieto[1,4], Andrea Vázquez-Ingelmo[1,3,4], Roberto Therón[1,2,3], Francisco J. García-Peñalvo[1,3,4] and Martín Martín-González[5]

[1]GRIAL Research Group. University of Salamanca
[2]VisUSAL Research Group. University of Salamanca
[3]Department of Computer Science. University of Salamanca
[4]Research Institute for Educational Sciences. University of Salamanca
[5]UNESCO Chair in University Management and Policy. Technical University of Madrid

{juancb, josecarlos.sp, andreavazquez, theron, fgarcia}@usal.es
martin.martin@upm.es

Abstract. This paper presents a research work that analyzes the effect of redirecting users between two different versions of a web form after they have started the questionnaire. In this case, we used a web form proposed by the Spanish Observatory for Employability and Employment (OEEU) that is designed to gather information from Spanish graduates. These two versions are different as follows: one of them is very simple and the other one includes several changes that appeared in the literature related to users' trust, usability/user experience and layout design. To test the effect of redirecting users between both versions of the web form, we used a group of users that already have started the questionnaire and redirect them to the other version; this is, we changed the web form version they use to the other version and measure how this change affects them. This experiment has shown some promising results, which lead to enhance and extend the experience to bigger populations and other kind of changes in the user interfaces.

Keywords: Human-Computer Interaction, HCI, Web forms, Online Questionnaire, User Experience, Performance.

1 Introduction

Currently, the web forms are one of the most used ways to get information from the users [1]. The easiness of deploying web forms in websites and the users' habit to use them have converted online questionnaires in a pervasive tool to gather information. Thus, the research on engaging users to fulfil questionnaires and web forms is a research area that evolves continuously and relates to other areas like user experience [2], psychology [3], data retrieval [4, 5], etc.

Regarding the user experience (UX), the work done to improve web forms in recent years has been carried to advance in issues like how to properly design the web layout [6, 7], how to design interesting user experiences [8-10], on how to communicate effectively with the user to improve the trust on the web form [3, 11], formalize usability standards [9, 12, 13], etc.

Following some of these research lines, we are working on how to propose and design different versions of the same web form to measure and detect which versions are the best regarding to improve the users' performance [14, 15]. From a methodological perspective, our approach is based on A/B tests. Following the A/B tests foundations, we show different variations of a website (in this case of a web form) randomly to different users and measuring what variation is the most effective (in terms of click-performance, task-performance, etc.). In our current research, we have developed two main versions: one based on extreme simplicity (with no visual effects or transitions, a simple layout, etc.) and other less simple that include characteristics like transitions, elements that enhance the user's trust on the website, a more elaborated layout, etc. In the following section (methodology) both versions will be explained in depth.

The different versions have been developed for a web form used by the Spanish Observatory for Employability and Employment (OEEU in its Spanish acronym) [16]. The web form is intended to gather data about how graduates get employment after they left the university. In this case, this online questionnaire is the most important tool for the Observatory to obtain data and information, without it the Observatory would not have data to develop their studies about employability and employment.

In the case of this research, we have applied the two different versions randomly to the graduates that participate in the OEEU's data gathering questionnaire. After a while, we began a reinforcement phase where the graduates that dropped out the questionnaire (or did not start it) would be given an opportunity to participate again. In this reinforcement phase, we redirected users between both versions (swapping users between the simpler and the less simple) depending on their performance, to test how varying the web form's features and complexity would affect the users' performance in completing the questionnaire.

So, this paper presents the results of a pilot study carried by the Spanish Observatory for Employability and Employment and the GRIAL Research Group at the University of Salamanca (Spain), with the objective of examining the effects of changing the web form layout and features on the finalization rate of the users that have already initiated the answering process.

The article is organized in three sections. The first one is dedicated to describing the methodology. It details the description of the different versions of the questionnaire, and the redirection process, as well as the research design, and the sample. After that, we present the results obtained, including the hypothesis testing using three-dimensional contingency tables. Finally, we will close the communication with a discussion and a brief series of conclusions.

2 Methodology

This section presents the methodology and other relevant aspects of this research.

2.1 Different versions of the web form

The two versions of the web form, as explained before, are differentiated basically because the first version is the simplest one (in the case of a web form) and the other is a bit more complex and has several features designed to engage user and develop an effective communication and relationship. The different changes proposed between both versions are based on different proposals retrieved from the literature and design guidelines, as explained in [14].

In the case of the simplest version, which we call "A" version, the web form is a basic form built using Bootstrap 3, with only one logotype (from OEEU), and a simple combination of visual elements with basic colors like white, blue and green (following the Bootstrap's style). A basic example of how is the layout of this "A" version can be found in this PDF (content in Spanish) or in [14].

In the case of the "B" version (the second one), it changes several things aimed at developing a closer relationship with the user (as proposed in the Social Exchange Theory [1]), enhancing the user's trust on the web form owner and its intentions [3, 11], improving issues related to user experience [17], usability [18] and interface design of the questionnaire [7, 19]. Specifically, the changes introduced in the "B" version of the web form were:

1. **Adequacy of the image to the other digital products of the Observatory**. In this change, related to enhance the users' trust, we planned to update the visual layout of the web form to meet the OEEU's design guidelines used in other of their digital products.
2. **Inclusion of the Observatory's logo and university's logo.** In this case, this is a change also related to building trust. It proposes to include the OEEU's logo in the web form header, as well as the logotype from the university where the student graduate.
3. **Inclusion of a progress bar in the questionnaire.** In this case, the proposed change was focused on improving the user experience with the web form. It is a simple change that consists (only) in including a progress bar that informs users about their progress in the task of finalizing the questionnaire.
4. **Present a visual focus animation on concrete actions.** Another proposal related to the usability and user experience. In this case, this change was designed to get the user attention and minimize the effort on using the web form. In this case, for example, the web form will auto scroll smoothly to the following question after the user responds to the previous one.
5. **Deactivation of control elements when an action is initiated.** This proposal consists on deactivating visual elements (like buttons) while they respond or complete an action requested by the user. For example, deactivating a button after the user clicks on it while the action triggered is completed.
6. **In related elements, instead of having smaller and more specific groupings, use some larger grouping, following the Gestalt principles on grouping.** This change was specially designed for large groups of questions/answers. Usually, in the web form, questions that include subquestions and nested response options are arranged in tables. For

example, following the proposal, the header of a table would be fixed while the content can be scrolled up and down. It seeks to ensure that the large dimensions of analysis in some points of the questionnaire are grouped to avoid user fatigue and reducing the users' cognitive load when dealing with large tables or complex visual elements.

To get more information or find visual examples of these changes, we refer the reader to [14].

2.2 Redirection process

As previously commented, the users were initially assigned randomly to use the "A" or "B" version of the web form. While users were using the web form, we analyzed what kind of factors (users' personal factors, technological aspects, etc.) were related to the users' performance in completing the questionnaire. By using predictive models and clustering techniques, we figured out the behaviors shared among users, what were their common characteristics, etc., to find patterns that define what lead users to achieve better performance metrics. As an example, in this previous research to know the most relevant user factors regarding to performance, in general we found that users have better average performance using the simplest version ("A"), except for those users that employ mobile or other devices with special specs like big screens or screens with an extremely good resolution, etc. Specifically, we found that users that meet the following criteria had better performance metrics in the "B" version:

1. Users that utilize Android devices with screens of 3- or 4-pixel ratio.
2. Users that accessed to the web form using large iPhone devices (iPhone 6 Plus, 6s Plus, or 7 Plus).
3. Users that use Android tablets.

So, in general, in the reinforcement phase, all users that randomly were assigned to use the "B" version and did not meet these conditions were redirected to the "A" version. On the other hand, all users that used the "A" version in the initial stage and meet those conditions (or rules) were redirected to the "B" version in the reinforcement phase. This kind of rules were used to change users between both versions.

2.3 Research design and Sample

The study presented in this paper is framed within another big (and more generalist) study about web forms and user experience. The whole experiment was conducted with more than 6700 users (graduates). Specifically, the questionnaires about employment were initiated by 6738 users, from which 5214 finalized the process (finalization ratio of 77.38%).

As previously commented, we ran a reinforcement phase after the first round of questionnaires; in this reinforcement stage were invited again to participate all those students that did not completed (or started) the web forms at the first round.

In the case of the study, 123 users were involved that participated in the first round and did not finalized the questionnaire, reentering again on the questionnaire during the reinforcement phase.

Analyzing these users, we studied the users' performance related to each version of the web form and we swapped users between both versions to test what is the effect of this change in their performance.

To do so, we proposed a quasi-experimental research design with a control group. Following this design, we divided the users in two groups: the experimental group (89), composed by the users that were redirected from one questionnaire design to a different one, and the control group (34) composed by those users that remained in the same questionnaire design.

After the application of the different treatment to each one of the groups we compared the differences in the finalization rate using three-dimensional-contingency tables and chi-squared to analyze the impact of increasing or decreasing the complexity of the questionnaire. In consequence, we pose the following hypotheses:

H1 The redirection to a different version of a questionnaire will have an impact on the finalization rate.

H2 The redirection of users from a text plain questionnaire to one with more complex elements will have an impact on the finalization rate.

H3 The redirection of users from a questionnaire with complex elements to a plain text one will have an impact on the finalization rate.

3 Results

As mentioned before, to assess the general effect of the redirection of the users on the improvement of the finalization rate we have used an approach based on the use of three-dimensional-contingency tables, a methodology of analysis useful to compare the effect of a variable in the relationship of the other two variables.

In this case, we will begin analyzing the effect of the web form version in the users' finalization rate to decide which the users will be redirected following our rules in the redirection phase. In other words, we analyze the effect of the redirection considering the version of the questionnaire to decide which users will be redirected during the reinforcement phase.

As a first step for the analysis, we elaborated the three-dimensional contingency table for this variable to see if there are observable differences at plain sight between the control and the experimental group considering the version of the questionnaire to which the users were redirected (Table 1).

Table 1. Three-dimensional contingency table for questionnaire redirection.

Version Redirected To	Group	Finalized	
		Yes	No
A	Control	10	4
	Experimental	38	33
	Total	48	37
B	Control	9	11
	Experimental	7	11
	Total	16	22

As we can see in Table 1, there were some minor differences in the finalization rates of the control and the experimental groups in both versions of the questionnaire. In consequence, we proceed with the calculation of the chi-squared index to find out wether there were any relation between being redirected and the finalization of the questionnaire (Table 2). The results lead to the rejection of the hypothesis H1 (*the redirection to a different version of a questionnaire will have an impact on the finalization rate*) in both versions for a significance level of 0.05.

Table 2. Results of chi-squared for questionnaire redirected to

Vertical Redirected To	Value	df	Significance
A	1.526	1	0.217
B	0.145	1	0.703

Lastly, to contrast the last two hypotheses, we perform the same procedure, but considering the complexity change. This way, we measured if the users changed, from a simpler version of a questionnaire to a more complex one or if at the contrary, the users change from a complex questionnaire to a simpler one. As in the previous case, we begin elaborating the three-dimensional contingency table (Table 3), but only with the users that were redirected from questionnaire A to questionnaire B or from questionnaire B to questionnaire A.

Table 3. Three-dimensional contingency table for questionnaire redirected from.

Version Redirected From	Group	Finalized	
		Yes	No
A to B	Control	10	4
	Experimental	1	8
	Total	11	12
B to A	Control	9	11
	Experimental	17	19
	Total	26	30

After analyzing the results, we performed the correlational analysis to know if there were any relation between being redirected and the finalization of the questionnaire. For the case of the change from questionnaire A to questionnaire B we used Fisher's exact test, due to the size of the groups (Table 4). The results support hypothesis H2 (*the redirection of users from a text plain questionnaire to one with more complex elements will have an impact on the finalization rate*), but reject hypothesis H3 (*the redirection of users from a questionnaire with complex elements to a plain text one will have an impact on the finalization rate*).

Table 4. Results of correlation for questionnaire redirected from.

Vertical Redirected To	Value	df	Significance
A to B	-	-	0.009
B to A	0.026	1	0.873

4 Discussion

The results obtained in the present research entail a series of implications both for theory and practice of the design of online questionnaires.

Firstly, we would like to highlight the rejection of the hypothesis H1. This can be caused by the fact that the redirection rules were based on the behavior of the users that entered the questionnaire for the first time, which indicates the need to deepen in the analysis of the behavior of the people that resume the questionnaire completion process to know how this variable may impact the finalization rate of this kind users.

In this line, the results of the analysis of the hypotheses 2 and 3 suggests that increasing the number of design elements in the questionnaire has a negative effect on the finalization rate, while the redirection to a plain text questionnaire does not have any effect.

A possible explanation may lay in the users' motivation. The users participating in this pilot study are those that have already tried to complete the questionnaire but abandoned the process, which make very likely that their motivation levels were low. Taking this into account, it is logical to think that when these users resume the process, finding a questionnaire with more design elements that the one that they initiated in the first place, may feel discouraged.

As a consequence, we believe that avoiding this kind of redirections in the case of users that are resumed the questionnaire is advisable.

Another possible explanation could be related to some of the elements introduced or removed when changing the web form version. In this sense, the experiment in its current setup cannot allow us to ascertain which changes affect more or less to user performance. For that reason, one of the main shifts and improvements in this experiment and research could be to divide more the versions, achieving the same number of versions than the total changes. So, in this case, we would be able to detect how affects each individual change to the user's performance compared to the simplest version.

5 Conclusions

This paper presents a novel research work that analyzes the effect of redirecting users between two different versions of a web form after they have started the questionnaire. We used a web form proposed by the Spanish Observatory for Employability and Employment (OEEU) that is designed to gather information from Spanish graduates. To test the effect of redirecting users between both versions of the web form, we used a group of users that already have started the questionnaire and redirect them to the other version; this is, we changed the web form version they use to the other vertical and measure how this change affects them.

In general, the results are quite promising and encourage us to continue the labor of researching how different changes in web forms affect users' performance. In this case, we can conclude that if we redirect users between two versions of a web form, the change will be negative if the user is redirected to a more complex version and will not have effect if it is redirected to a simpler version. In the future, we would like to enhance and extend the experiment to bigger populations and other kind of changes in the user interfaces to verify these initial results. Also, we would like to test how users feel the change and what is their opinion about the change (to compare also their feelings and perception to their performance).

Acknowlegments. The research leading to these results has received funding from "la Caixa" Foundation. The author Juan Cruz-Benito would like to thank the European Social Fund and the *Consejería de Educación* of the *Junta de Castilla y León* (Spain) for funding his predoctoral fellow contract. This research has been cofounded by the University of Salamanca through the program of financial aid for the predoctoral contract of José Carlos Sánchez-Prieto (*Programa III: Ayudas para contratos Predoctorales*) cofounded by Banco Santander. This work has been partially funded by the Spanish Government Ministry of Economy and Competitiveness throughout the DEFINES project (Ref. TIN2016-80172-R).

References

1. Dillman, D.A.: Mail and internet surveys: The tailored design method. Wiley New York (2000)
2. Jarrett, C., Gaffney, G.: Forms that work: Designing Web forms for usability. Morgan Kaufmann (2009)
3. Seckler, M., Heinz, S., Forde, S., Tuch, A.N., Opwis, K.: Trust and distrust on the web: User experiences and website characteristics. Computers in Human Behavior 45, 39-50 (2015)
4. Stieger, S., Reips, U.-D.: What are participants doing while filling in an online questionnaire: A paradata collection tool and an empirical study. Computers in Human Behavior 26, 1488-1495 (2010)
5. Kveton, P., Jelínek, M., Klimusová, H., Voboril, D.: Data collection on the Internet: Evaluation of Web-based questionnaires. Studia psychologica 49, 81 (2007)
6. Flavián, C., Gurrea, R., Orús, C.: The effect of product presentation mode on the perceived content and continent quality of web sites. Online Information Review 33, 1103-1128 (2009)
7. Seckler, M., Tuch, A.N., Opwis, K., Bargas-Avila, J.A.: User-friendly locations of error messages in web forms: Put them on the right side of the erroneous input field. Interacting with Computers 24, 107-118 (2012)
8. Bargas-Avila, J.A., Brenzikofer, O., Tuch, A.N., Roth, S.P., Opwis, K.: Working towards usable forms on the worldwide web: optimizing multiple selection interface elements. Advances in Human-Computer Interaction 2011, 4 (2011)
9. Bargas-Avila, J., Brenzikofer, O.: Simple but crucial user interfaces in the world wide web: introducing 20 guidelines for usable web form design (2011)
10. Anderson, S.P.: Seductive Interaction Design: Creating Playful, Fun, and Effective User Experiences, Portable Document. Pearson Education (2011)

11. Shneiderman, B.: Designing trust into online experiences. Communications of the ACM 43, 57-59 (2000)
12. ISO/DIS: Draft BS ENISO 9241-220 Ergonomics of human-computer interaction. Part 220: Processes for enabling, executing and assessing human-centred design within organizations. (2016)
13. ISO/DIS: Ergonomics of human-system interaction — Part 125: Guidance on visual presentation of Information. (2013)
14. Cruz-Benito, J., Therón, R., García-Peñalvo, F.J., Sánchez-Prieto, J.C., Vázquez-Ingelmo, A., Martín-González, M., Martínez, J.M.: Improving Success/Completion Ratio in Large Surveys: A Proposal Based on Usability and Engagement. In: Zaphiris, P., Ioannou, A. (eds.) Learning and Collaboration Technologies. Technology in Education: 4th International Conference, LCT 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part II, pp. 352-370. Springer International Publishing, Cham (2017)
15. Cruz-Benito, J., Vázquez-Ingelmo, A., Sánchez-Prieto, J.C., Therón, R., García-Peñalvo, F.J., Martín-González, M.: Enabling Adaptability in Web Forms Based on User Characteristics Detection Through A/B Testing and Machine Learning. IEEE Access 5, (2017)
16. Michavila, F., Martínez, J.M., Martín-González, M., García-Peñalvo, F.J., Cruz-Benito, J.: Barómetro de Empleabilidad y Empleo de los Universitarios en España, 2015 (Primer informe de resultados). (2016)
17. Recabarren, M., Nussbaum, M.: Exploring the feasibility of web form adaptation to users' cultural dimension scores. User Modeling and User-Adapted Interaction 20, 87-108 (2010)
18. Huber, W., Vitouch, P.: Usability and accessibility on the internet: Effects of accessible web design on usability. In: International Conference on Computers for Handicapped Persons, pp. 482-489. Springer, (2008)
19. Pengnate, S.F., Sarathy, R.: An experimental investigation of the influence of website emotional design features on trust in unfamiliar online vendors. Computers in Human Behavior 67, 49-60 (2017)

6.12 Appendix L. Proposing a machine learning approach to analyze and predict employment and its factors

Article in Press IJIMAI journal

Proposing a Machine Learning Approach to Analyze and Predict Employment and its Factors

Francisco J. García-Peñalvo[1,3,4]*, Juan Cruz-Benito[1,3,4], Martín Martín-González[5], Andrea Vázquez-Ingelmo[1,3,4], José Carlos Sánchez-Prieto[1,4], Roberto Therón[1,2,3]

[1] GRIAL Research Group. University of Salamanca (Spain)
[2] VisUSAL Research Group. University of Salamanca (Spain)
[3] Department of Computer Science. University of Salamanca (Spain)
[4] Research Institute for Educational Sciences. University of Salamanca (Spain)
[5] UNESCO Chair in University Management and Policy. Technical University of Madrid (Spain)

Received 3 November 2017 | Accepted 22 January 2018 | Published 2 February 2018

Abstract

This paper presents an original study with the aim of propose and test a machine learning approach to research about employability and employment. To understand how the graduates get employed, researchers propose to build predictive models using machine learning algorithms, extracting after that the most relevant factors that describe the model and employing further analysis techniques like clustering to get deeper insights. To test the proposal, is presented a case study that involves data from the Spanish Observatory for Employability and Employment (OEEU). Using data from this project (information about 3000 students), has been built predictive models that define how these students get a job after finalizing their degrees. The results obtained in this case study are very promising, and encourage authors to refine the process and validate it in further research.

Keywords

Employability, Employment, Artificial Intelligence, Machine Learning, Random Forest, Academic Analytics, OEEU.

DOI: 10.9781/ijimai.2018.02.002

I. Introduction

The concept of employability has steadily gained importance in recent years, becoming one of the pillars of the European educational strategy within the European Higher Education Area (EHEA) framework. However, the empirical research is still insufficient to build a strong theoretical foundation. It is worth noting that applied research on employability has, nowadays, an exploratory approach. This is because this research area presents difficulties regarding to have adequate, reliable and updated data, as well as this early status not only prevents agreement on research results are reached, but also poses many questions as to which methodologies and approaches are most appropriate to address these issues. For these reasons, the area is still growing and need to push the outcomes to further research levels.

* Corresponding author.

E-mail address: fgarcia@usal.es

Several research projects have been developed in recent years to provide more information on the employability of graduates, many of which have been driven by the OECD and the European Commission. These projects have faced at least two problems: first, the lack of a single, consensual definition of employability and, secondly, the difficulty of obtaining summary indicators to assess it.

Indeed, employability is a theoretical construct whose definition varies according to academic discipline and the perspective used, as well as the socioeconomic context to which it refers. There is no clear consensus on the factors that compose or determine it, nor on the employment outcomes to which it leads. Therefore, evaluating employability is a tough task. In any case, given the complexity of the notion of employability, it would be worth using several variables and indicators that assess different labor, educational and sociodemographic issues, rather than a summary indicator.

Most of the studies on employability that have been developed since the 1990s have focuses on identifying the competencies that graduates will need throughout their career path. Some have gone further,

Please cite this article in press as:
F. J. García-Peñalvo, J. Cruz-Benito, M. Martín-González, A. Vázquez-Ingelmo, J. C. Sánchez-Prieto, R. Therón. Proposing a Machine Learning Approach to Analyze and Predict Employment and its Factors, International Journal of Interactive Multimedia and Artificial Intelligence, (2018), http://dx.doi.org/10.9781/ijimai.2018.02.002

introducing other variables related to the training and education that offer the academic institutions, the sociodemographic context, the institutional and normative framework and the productive structure ("broad" approach as presented in [1]).

This kind of research, despite its initial status, is focused on develop successful strategies and outcomes that could help policymakers and institutions to enhance and promote those detected factors that contribute to get more chances of employment and better employments. For that reasons and its application in the society, it is possible to affirm that the project is in the scope of emergent areas like the Academic Analytics [2-6] or Institutional Intelligence [7, 8].

This paper aims to present a new method to analyze employability factors and to analyze how people gets employed. To achieve that, this paper proposes a machine-learning-based approach that produce predictive models on employment, providing the main factors that affect the predictive model and finding the most relevant ones. This approach contrasts to the previous state of the art in this research area. As will be explained in the Background section, previous approaches are based on basic statistical processes and tries to accomplish the problem of employment and its factors as a whole, instead of weighting the relevance of each factor to build more complex models. To illustrate these considerations, the paper provides a case of study where has been applied the approach and shows some promising results.

The research presented in this paper is developed under the scope of the Spanish Observatory for University Employability and Employment (OEEU in its Spanish acronym) [9]. This observatory gathers data about employment and employability parameters among the Spanish graduates (after they leave the university) to analyze the information they provide and understand what the employment trends and most important employability factors are for this population [10]. To accomplish this mission, the observatory has developed a complex information system [6, 11-13] that collects and analyzes data to present the insights to the researchers [14, 15]. These data collected are used as a dataset to test the machine learning approach that will be presented in the following sections.

This paper has the following structure: second section (Background) presents the state of the art in the case of research applied on employability and employment analysis. Third section (Proposal) describes the machine learning approach and the methods and materials used in this research. Section fourth (Case Study: OEEU) shows the case study developed and the initial results achieved. Fifth section (Discussion) discusses about the implications of the research presented and the results achieved. Sixth section (Conclusions) finalizes the paper with some final remarks and introduces some future work.

II. Background

One of the main competencies studies promoted by the OECD was the "Definition and Selection of Key Competences" (DeSeCo). Since the first editions of the Programme for International Student Assessment (PISA) it had become clear that job success depended on a much greater range of competencies than those considered in the project. The DeSeCo project was created to identify these key competencies, aiming to serve as a framework to guide and complement two international programs to evaluate competencies: the aforementioned PISA and the Adult Literacy and Lifeskills (ALL). The DeSeCo project began in 1997 and ended in 2003, when it published the final report entitled "Key Competencies for a Successful Life and Well-functioning Society" [16]. In this project worked academics and experts from different fields (sociologists, philosophers, psychologists, economists, anthropologists, historians, statisticians, educators, etcetera) and social institutions (political parties, unions, employers, associations, etcetera) to define and figure out key competencies based on previous research. The final list of competencies was discussed in depth in two international symposia until achieve an agreement on the most important. This project is one of the foundational approaches and projects for the employment and employability analytics knowledge area.

One of the most influential projects in the EHEA when defining, identifying and classifying competences was the Tuning Educational Structures in Europe (usually called Tuning project). This project was funded by the European Commission within the Socrates framework. The project was divided into two phases, the first of which was more significant. It was active between 2000 and 2002. Its main objective was to figure out and classify the competences that the graduates require on their career path. Experts from different fields of knowledge (Business Administration, Geology, History, Mathematics, Physics, Education and Chemistry), from several European countries (Germany, Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Iceland, Italy, Norway, Netherlands, Portugal, Spain and Sweden) worked on the project. A total of 5803 graduates, 944 employers, and 998 scholars were surveyed; with the participation of more than 100 educational institutions of the European Union. The research finally divided the competencies into two main groups — specific and generic— and, in turn, the generic ones into three groups: instrumental, interpersonal and systemic.

Driven by the creation of the EHEA, different employability and competency projects have been developed that have enabled comparisons between European countries and universities using a common methodology. One of these projects was the "Higher Education and Graduate Employment in Europe" (usually called CHEERS project — "Careers after Higher Graduation. A European Research Study"). The project was promoted and financed by the European Commission within the "Targeted Socio-Economic Research Programme" (TSER). It began in 1997 and ended in 2000. It was led by The International Centre for Higher Education Research at the University of Kassel (Germany) and included other countries like Germany, Austria, Spain, Finland, France, Iceland, Italy, Japan, Norway, Netherlands, United Kingdom, Czech Republic and Sweden. Between 1998 and 2000 the different research groups sent a standardized questionnaire to the graduates who had completed their studies in the academic year 1994/1995. 37000 subjects were interviewed (about 3000 from each university). The information collected was based on their studies and the career path to analyze the relationship between higher education and employment (job position, mismatch in the labor market, etcetera). An important part of the questionnaire, in which the project had put special emphasis, was the assessment of the level of graduates' competencies and the level that they required by employers.

The CHEERS project was the starting point for the development, in 2006, of "The Flexible Professional in the Knowledge Society: New Demands on Higher Education in Europe" project, usually called REFLEX project. This project aimed to answer, among others, the following questions: what competences do graduates require fulfilling the demands from the modern knowledge society? To what extent has higher education provided these competencies? How can the mismatches between acquired and required competencies be solved?; To what extent are the graduates' expectations met? [17]. It was funded by European Union within the 6th Framework Programme (FP) for Research and Technological Development. It was led by the Research Center for Education and the Labor Market of the University of Maastricht. 14 European countries (Germany, Austria, Spain, Finland, France, Italy, Norway, Netherlands, United Kingdom, Belgium, Czech Republic, Portugal, Switzerland, Estonia) and Japan participated. The methodology and the questionnaire were like that adopted in the CHEERS project. A total of 40787 graduates were surveyed in 1999/2000, 5500 of which corresponded to graduates in

Spanish universities (National Agency for Quality Assessment and Accreditation of Spain, 2008). Once again, information about the competencies that the project researchers considered relevant for the promotion of employability was compiled.

In Spain, the study carried out by the Catalan University Quality Assurance Agency (AQU) is one on the most important at the regional level. AQU conducts a telephone survey periodically (every three years) since 2001. The sample is made up of university graduates who finished their studies in any Catalan university three years before the date on which they were surveyed. Among other aspects the outcomes of this study are reports that analyze information related to the quality of employment, job stability, earnings, education-job skills match, job satisfaction, the process of finding a job, mobility, students' satisfaction with their studies, etc. In relation to competences, the questionnaire incorporates a section about skills acquired and their usefulness in the workplace.

Among the latest initiatives to evaluate the competencies of graduates in Spain highlight the *Libro Verde sobre la empleabilidad de los egresados de la Comunidad Valenciana* (Green book on the graduates' employability of Valencian Community). This book was published in 2013 because of a project promoted and funded by the *Generalitat Valenciana* and developed by the Valencian Agency of Assessment and Prospective (AVAP) and the universities of the region. The report presents the results of a study aimed to clarify the employability and employment situation of graduates at the *Comunidad Valenciana (*Valencian Community) and provided a series of conclusions and recommendations to reinforce and improve their employability. In this project was carried out a survey for those who had completed their short and long cycle studies during the years 2008 and 2009, comprising 2099 graduates in the study. This survey gathered information about the level of mastery of 21 generic competences that the graduates should have, as well as the level required by employers and the level acquired in universities.

Regarding to the methodologies applied in these studies, it is worth noting that they mainly use descriptive statistics, basically applying basic measures of central tendency and frequency distribution. They also rely mainly on basic charts to present the information in a simple way, basically using of histograms and bar charts. As an example, it is common to find in this type of studies the percentage distribution of graduates according to the strategy used in the search for employment, the time (months) taken for graduates to find their first job, the average earnings, the percentage of graduates satisfied with their jobs or their studies, the distribution of the graduates according to the education and job match, the average level of competences required by employers and acquired at the universities, etc. The number of variables involved in these studies varies notably depending on the scope of the project and its objectives. Among the largest is the REFLEX project, which almost included 500 variables distributed in 11 categories: study program, other educational and related experiences, transition from study to work, first job after graduation, employment history and current situation, current work, work organization, competencies, evaluation of study program, values and orientations, about yourself, plus the variables of country and study identification.

This kind of studies have attracted the attention from numerous researchers from different disciplines who, given the increasing availability of data, have been able to carry out some empirical papers. Thus, much of the research has focused on identifying the competencies required by employers and evaluating their impact on employment [18-21] as well as assessing the mismatches between acquired and required competencies [22, 23]. In addition, these research works allowed deepening the relationship education-labor market, as well as the mismatches and their effects [23-28]. They have applied all kinds of analytical methodologies. Although there is no common practice, they use more sophisticated techniques and tools than those applied by general studies, linked to econometrics, psychometry and other quantitative and qualitative methods from research in social sciences.

III. Proposal

In the case of this paper, researchers followed this kind of new approaches. Using novel methods and techniques like machine learning could open new possibilities and ways to work in assessing employability and employment. Also, these methods could unleash new ways for manage huge amounts of information as a whole but considering each factor in a weighted way within the predictive models to be built. Previously some projects use basic statistics due the difficulty of handling big datasets in a non-automated mode, but in our way, the same procedure can be used to crunch all data related jointly.

In general, the proposal for the analysis (based on machine learning) follows common principles in data science regarding data structuration, tidy data approaches, etc. [29-31]. Because it is a proposal and it is on its initial stage, the approach should be explainable to assess its appropriateness. For this reason, the machine-learning process has been implemented in a white-box way; thus, the researchers have selected algorithms and methods to make the workflow explainable.

Moreover, these main principles, the different details for the analysis pipeline, and methods used in this research are presented below. All of these details are explained in the following workflow (available at [32]):

1. Retrieve dataset about students from OEEU's information system.
2. Filter the desired fields from the datasets and enclosed them in data frame (a data structure like a table).
3. Data cleaning: remove noise data, remove columns (variables) with too many null (NaN) values, and remove all students who have only partial.
4. Normalize data with the One-hot encoding algorithm for categorical values in columns [33].
5. Considering the data gathered and the kind of variable (labeled) to predict (student gets employed or not), the algorithm to use must be related to supervised learning. This is because this kind of algorithm makes predictions based on a set of examples (that consist of a labeled training data set and the desired output variable). Moreover, regarding the dichotomous (categorical) character of the variable to predict, the supervised learning algorithm to apply must be based on classification (binary classification, as we have a label of finalization equal to true or false). According to the authors' previous experience, the possibility of explaining results and the accuracy desired for the classification, a Random Forest classifier algorithm [34] was selected. In this step, the Random Forest algorithm was executed repeatedly to determine the best setup for the dataset given (obtaining the most adequate parameters for the execution).
6. With the best configuration found, train the random forest algorithm (with 33.33% of the dataset) and obtain the predictive model.
7. Using the predictive model, obtain the most important features for the predictive model.

At this point, researchers have built a model that could predict if a person will get employed or not, showing also what are the factors that affect more the result. After that, researchers could use these factors to filter information, generalize knowledge across the dataset, extract what values on these factors lead to get employed or not, etc. As an example, using these most relevant factors, researchers could clusterize students to gain deeper knowledge about what are the main characteristics between those who get an employment and those who do not.

In general, in this paper will be demonstrated how this kind of approach could be suitable for the goal of modelling employment and its factors. To do that, the algorithms and the code used is available publicly at [32]. Unfortunately, due the privacy restrictions that affect the OEEU will not be revealed some kind of data involving individual graduates, universities or other sensitive information, it will be only displayed aggregated (and anonymized) information. For that reason, regarding the processes related to analyze the factors that define the model, only will be shown some generalist figures that could illustrate the procedure and give some clues about a real implementation.

The programming language used to conduct all the analyses and calculations was Python. The Python software tools and libraries used to code and test the approach were:

- Pandas software library [29, 35], to manage data structures and support analysis tasks.
- Scikit-learn [36] library, to accomplish the machine learning workflow [33].
- Jupyter notebooks [37-39], to develop the Python code used in this research.

IV. Case Study: OEEU

As presented in the previous sections, the projects that investigated employability and employment disciplines varied in number and type of variables available, scope of the project, etc. For this paper, researchers used the information gathered by the Spanish Observatory for Employability and Employment (OEEU). This project keeps information about 182000 students graduated from degree and master studies. It includes about 400-500 variables per each edition of the study (one edition about degree graduates – 134129 students involved – and other edition about master studies –47822 students–).

The data used in this case study correspond to the information available from graduated students that finalized a degree in the course 2009-2010. This is, information about 134129 students, with 493 variables per student [9, 32]. Despite of the dimensions of the dataset, it is worth noting that not all the students have information for all the possible variables. The Observatory gathers the information by using two input methods: the raw records from the Spanish universities and the information provided by the students through fulfilling questionnaires. These two main sources of information have only part of the variables marked as required, for that reason some of them appear with empty values. Also, the fact of ingesting information via web forms (like in this case) make extremely difficult to get all the information, because the graduate can quit the web form in any moment. For that reason, are required methods that clean and wrangle the information like those presented in the previous section.

Apart of cleaning and wrangling the data properly, researchers have excluded some variables included in the OEEU's dataset, since they are related exclusively to some universities (the universities could add some questions to the OEEU questionnaire), and using only the common variables to all students in their test to create the predictive model related to employment. This reduced the dataset from 493 total variables to 383 possible variables per student. These 383 variables per student can be observed in the 8th cell at the provided notebook [32].

Following the workflow outlined in the previous section, after filtering the desired variables, researchers cleaned all those variables which presents to much *NaN* (empty) values. In this case, the threshold used to remove all the weak variables was 10%. This is, all the variables with more than 10% of empty values were discarded for the model construction. This threshold is strict to obtain a stronger model. There are other common procedures to deal with void variables or measures (fill the empty values with others from the dataset, with the mean of the column, etc.), but in this case, researchers preferred to avoid any kind of artificial data that could contaminate the result. After removing all these non-valuable variables, researchers dropped all the students that had any empty value in their information (completing by this way the data-cleaning stage). After this hard-cleaning process, researchers counted with 26 data variables from 9744 graduates. As previously commented, following other conservative methods to deal with empty values would lead to use more variables (columns) and observations (rows), but this is not the focus in this initial test of the approach presented.

After all this work in data preparation, began the machine learning phase. In this case, the third part of the dataset (third part of graduates) was marked as the portion to train the random forest algorithm. Also, researchers selected the variable to predict using the others. In this case, the variable to predict was *'haEstadoDesempleado'* (Have the student been unemployed?) which contains two possible values: 0 (false) value for those students that got a job after finalizing the degree, and 1 (true) for those students that were not employed.

After that, and with this 33% of the observations (3305 students) and the variable to predict, researchers tested programmatically the best setup for the random forest algorithm. This is, the best configuration values for the parameters *randomforestclassifier__max_features* and *randomforestclassifier__min_samples_leaf.* Using the best values found for the parameters, researchers executed (trained) the algorithm to get the corresponding predictive value.

Table I presents the quality metrics [40] of the predictive model built to predict the graduates' employment. As displayed, the precision of the predictive model classifying and predicting the employment or not was of 0.71 (where 0 is the worst precision and 1 the best).

TABLE I. Results of the Predictive Model Built

	Precision[a]	Recall[b]	F1-score[c]	Suport[d]
False	0.73	0.12	0.20	1066
True	0.70	0.98	0.82	2239
Avg / total	0.71	0.70	0.62	3305

[a]The precision is the ratio *tp / (tp + fp)* where *tp* is the number of true positives and *fp* the number of false positives. The precision is intuitively the classifier's ability of not labeling as positive a sample that is negative. This score reaches its best value at 1 and worst score at 0.

[b]The recall is the ratio *tp / (tp + fn)* where *tp* is the number of true positives and *fn* the number of false negatives. The recall is intuitively the ability of the classifier to find all the positive samples. This score reaches its best value at 1 and its worst score at 0.

[c]The F1 score can be interpreted as a weighted average of the precision and recall, where an F1 score reaches its best value at 1 and its worst score at 0. The relative contribution of precision and recall to the F1 score is equal. This score reaches its best value at 1 and its worst score at 0.

[d]The support is the number of occurrences of each class in each predicted label.

On the other hand, the crosstab that expresses the number of good and bad predictions for the predictive model can be found in Table II.

TABLE II. Crosstab for the Predictive Model Built

	False (predictions)	True (predictions)
False (actual)	125	941
True (actual)	46	2193

Following the process, the important factors found for the model are presented in the Table III. Despite the random forest provide a importance score for all the variables involved in the predictive model, researchers stated 0.05 as the minimum value to consider the factor as relevant. This is because 0.05 is a common value to ensure reliable results in several analytical processes.

TABLE III. Most Relevant Factors for the Predictive Model Built

Name of the variable	Explanation	Importance score
universidad_id	The university where studied the graduate	0.437347
otrosCriterios SeleccionarPuesto Trabajo_circunstancias Personal	The graduate's opinion about the relevance of choose a job depending on the conditions related to personal context: family conciliation, etc.	0.150274
sexo_id	Graduate's gender	0.107880
residenciaExtranjero Motivos_cod	Graduate's reasons to live abroad during the degree	0.097754
otrosCriterios SeleccionarPuesto Trabajo_prestigio	The graduate's opinion about the relevance of choose a job depending on the conditions related to the prestige of the employer, the tasks to be done or to the position within the company.	0.075377
titulacion_id	The degree studied by the graduate.	0.054354

The importance score varies between 0–1, where 1 is the best score and 0 the worst one

As previously presented, after obtaining the predictive model, researchers can use the most relevant factors to analyze in deep what are the specific situations (values of factors) that lead students to get or not a job. For example, using the most relevant factors, could be generated clusters that group graduates using their similar characteristics. As an example, Fig. 1 presents the different clusters obtained after applying a hierarchical clustering algorithm to the clean dataset using the factors as variables to group the users. The representation is truncated to show the more related clusters jointly (showing only 12), but in fact, applying the hierarchical clustering were obtained 55 different groups of students.

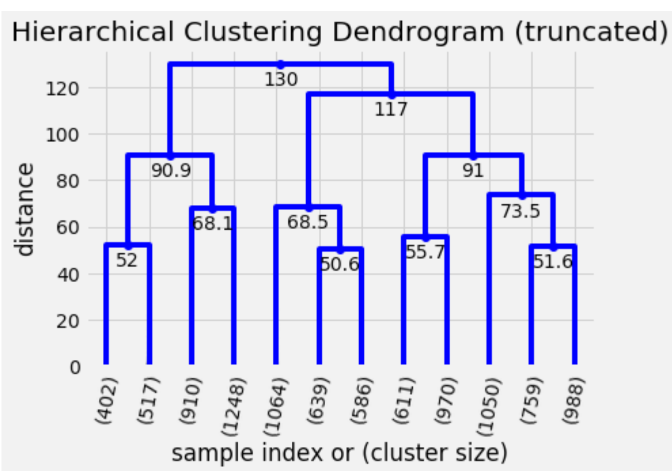


Fig. 1. Dendrogram that represents the clusters of graduates regarding to the most relevant factors detected in the random forest. Each leaf represents a different cluster obtained. The different values that appear near the claves display the Euclidean distance that explains the separation between the different clusters. Finally, the numbers below the leaves (at the bottom of the figure) present the number of users included in the corresponding cluster.

For example, if researchers choose one of the 55 resulting clusters, could observe that (cluster 22):

"229 students from 14 different universities, all of them women who studied one of 20 selected degrees, who do not consider the prestige of employer as a key factor to select a job, who lived abroad during the degree mainly because they work abroad, and do not consider the conditions related to personal context as a key factor to select a job have a chance of get job of 86,90%."

As previously explained, the full information about clusters and results obtained after applying the process is not fully provided in the paper or in an external notebook because of the Observatory's legal and privacy restrictions. Also, this kind of information is out of the scope of this paper, since it is focused on explaining how machine learning methods could be applied to this problem.

V. Discussion

As outlined in the background, the research on employability and employment is a knowledge area in development. The main projects developed in the previous years have pursued to define the main factors that define the employability and employment. The research methods used previously were related to basic statistics and simple analysis, despite some independent researchers went deeper in the methods, applying other related to econometrics, psychometry and other quantitative and qualitative methods of social research. One of the problems observed by the researchers when working with this kind of projects (i.e. in the case of the OEEU), is how to manage to handle large amounts of data and make more specific analysis. The approach presented in this paper regarding the application of machine learning methods, allows to automate some part of the analysis while it allows to gain general knowledge and deal with the problem of analyze how people gets employed from a broader perspective, understanding all the data as a whole.

The machine learning approach has been applied successfully previously by the authors and other researchers in fields like Human-Computer Interaction [41], education [42], etc. For that reason, was tried this approach in a complex research like employment and employability.

Considering the case study presented, the results are quite promising. Despite of the complexity of the data, and the different issues with the dataset, researchers have been able to get a predictive model with a 0.71 precision score (0 the worst score, 1 the best one). This result opens the possibility of keep working in this approach to enhance the results and try deeper analysis. Regarding other scores achieved like, F1, recall, etc., should be outlined that the model built performs poorly in detecting the real "False" values (not employed students), so it could lead to bad predictions regarding students without employment.

In this case, researchers are confident in that managing better the empty values and applying other kind of strategies in data cleaning and wrangling will allow to get better predictive models and outcomes. Following with the results and the case study, it is worth noting that the predictive model involved a support of 3305 observations (graduates), which is a high number in the case of a test like this. It is because researchers want to try a real case to validate the seminal idea of applying machine learning. Despite the model and procedures should be validated in a better way and tested more, the results achieved in a real case like this are considerable.

About the predictive model built, it is not surprising that the some of the most important factors that define if a student gets employment or not after graduating are those related to the university or the degree. Currently there are severe gaps regarding the employment ratio depending on the studies and the students' knowledge area. Also, the predictive model highlights other factor sadly known nowadays: the gender. There are many international studies that deals with the fact that the gender (specially for women) is a handicap to apply for some jobs and positions. The predictive model built found also this variable as a fundamental factor that define how graduates achieve an employment. On the other hand, the predictive model also presents other factors that are not as well-known as the previous ones: the reasons to live abroad during the studies, the importance of employer prestige to choose a job, or the facilities provided by the employer to adjust and balance personal context and life with the work. In general, it is clear that these factors are truly relevant in the model. These 5 factors sum a score of more than 0.8 out of the maximum 1 score that could be achieved by all the 383 factors included in the predictive model. Also, it is possible to think that the least relevant factor of these 5 (the degree obtained by the student) has a very low score (0. 054354), but this score is the weight of a solely factor in a model with 383 different factors, so achieve a 0.05 score out of 1 maximum score it is not low with this amount of different variables observed.

This kind of algorithmic approaches that include all the factors as part of possible models, shed light over some aspects avoiding previous bias and prejudices. In contrast to the background, where the international research community define the factors to study in each project, in this case, authors propose to use all the factors letting the machine learning algorithms to select on their own those truly relevant factors. This switch the *traditional* approach, making the exploratory process to depend only on the dataset available and adapting the focus to the facts and metrics obtained previously.

Also, are provided within the case study some examples on how to obtain deeper information and insights about the concrete metrics that affect the employment. This is, considering the most relevant factors, how to get common characteristics between those students who get an employment and those who do not. There are other many approaches to make this process possible, but the proposed using clustering analysis is evident and easy to understand and apply.

Regarding the implications of this research, authors agree to point out that this kind of approaches could open new possibilities on using data to enhance the education and students' opportunities in labor market. Following with the idea of Academic Analytics, this kind of employment analytics could be included in academic intelligence processes in universities and other higher education institutions. Also, policymakers could follow this kind of data-driven employability and employment analytics to design and propose new ways of preparing students to their professional future.

VI. Conclusion and Future Work

This paper presents a novel study in the field of employability and employment analytics. The main results achieved have been quite promising and encourage authors to continue the labor of improving the generation of predictive models for employability and employment. The nature of this kind of problems is extremely complex and varies on the time, but with this kind of algorithmic and automated processes could address it better than the traditional approaches. Based on the results, the authors are committed to continue developing the approach to get better results and improve the process until it could be applied successfully in further research works.

Acknowledgment

The research leading to these results has received funding from "la Caixa" Foundation. The author Juan Cruz-Benito would like to thank the European Social Fund and the *Consejería de Educación* of the *Junta de Castilla y León* (Spain) for funding his predoctoral fellow contract. This research has been cofounded by the University of Salamanca through the program of financial aid for the predoctoral contract of José Carlos Sánchez-Prieto (*Programa III: Ayudas para contratos Predoctorales*) cofounded by Banco Santander. This work has been partially funded by the Spanish Government Ministry of Economy and Competitiveness throughout the DEFINES project (Ref. TIN2016-80172-R).

References

[1] R. W. McQuaid and C. Lindsay, "The concept of employability," *Urban studies,* vol. 42, no. 2, pp. 197-219, 2005.

[2] P. Baepler and C. J. Murdoch, "Academic analytics and data mining in higher education," *International Journal for the Scholarship of Teaching and Learning,* vol. 4, no. 2, p. 17, 2010.

[3] J. Bichsel, *Analytics in higher education: Benefits, barriers, progress, and recommendations*. EDUCAUSE Center for Applied Research, 2012.

[4] J. P. Campbell, P. B. DeBlois, and D. G. Oblinger, "Academic analytics: A new tool for a new era," *EDUCAUSE review,* vol. 42, no. 4, p. 40, 2007.

[5] D. A. Gómez Aguilar, F. J. García-Peñalvo, and R. Therón, "Analítica Visual en eLearning," *El Profesional de la Información,* vol. 23, no. 3, pp. 233-242, 2014.

[6] F. Michavila, M. Martín-González, J. M. Martínez, F. J. García-Peñalvo, and J. Cruz-Benito, "Analyzing the employability and employment factors of graduate students in Spain: The OEEU Information System," presented at the Third International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'15), Porto, Portugal, 2015.

[7] Oficina de Cooperación Universitaria. OCU, "Libro Blanco Inteligencia Institucional en Universidades," ed: Madrid: OCU, 2013.

[8] F. J. García-Peñalvo, "Inteligencia Institucional para la Mejora de los Procesos de Enseñanza-Aprendizaje," *GRIAL Research Group*.

[9] F. Michavila, J. M. Martínez, M. Martín-González, F. J. García-Peñalvo, and J. Cruz-Benito, "Barómetro de Empleabilidad y Empleo de los Universitarios en España, 2015 (Primer informe de resultados)," 2016.

[10] W. Greller and H. Drachsler, "Translating learning into numbers: A generic framework for learning analytics," *Journal of Educational Technology & Society,* vol. 15, no. 3, pp. 42-57, 2012.

[11] A. Vázquez-Ingelmo, J. Cruz-Benito, and F. J. García-Peñalvo, "Improving the OEEU's data-driven technological ecosystem's interoperability with GraphQL," presented at the Fifth International Conference Technological Ecosystems for Enhancing Multiculturality 2017 (TEEM'17), Cádiz, Spain, October 18-20, 2017, 2017.

[12] A. Vázquez-Ingelmo, J. Cruz-Benito, F. J. García-Peñalvo, and M. Martín-González, "Scaffolding the OEEU's Data-Driven Ecosystem to Analyze the Employability of Spanish Graduates," in *Global Implications of Emerging Technology Trends*, F. J. García-Peñalvo, Ed. Hershey, PA: IGI Global, 2018, pp. 236-255.

[13] A. García-Holgado, J. Cruz-Benito, and F. J. García-Peñalvo, "Analysis of Knowledge Management Experiences in Spanish Public Administration," presented at the Third International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'15), Porto, Portugal, 7-9, October, 2015.

[14] J. Alcolea Picazo and S. Pavón de Paula, "Los datos como recurso estratégico Libro Blanco Inteligencia Institucional en Universidades (pp. 17-43)," *Madrid, Spain: OCU (Oficina de Cooperación Universitaria),* 2013.

[15] M. Zeleny, "Management support systems: Towards integrated knowledge," *Human systems management,* vol. 7, pp. 59-70, 1987.

[16] D. S. Rychen and L. H. Salganik, *Key competencies for a successful life and well-functioning society*. Hogrefe Publishing, 2003.

[17] National Agency for Quality Assessment and Accreditation of Spain, *Titulados universitarios y mercado laboral, Proyecto REFLEX*. Madrid: ANECA, 2008.

[18] R. G. Biesma, M. Pavlova, G. Van Merode, and W. Groot, "Using conjoint analysis to estimate employers preferences for key competencies of master level Dutch graduates entering the public health field," *Economics of Education Review,* vol. 26, no. 3, pp. 375-386, 2007.

[19] A. García-Aracil and R. Van der Velden, "Competencies for young European higher education graduates: labor market mismatches and their

payoffs," *Higher Education,* vol. 55, no. 2, pp. 219-239, 2008.

[20] H. Heijke, C. Meng, and G. Ramaekers, "An investigation into the role of human capital competences and their pay-off," *International Journal of Manpower,* vol. 24, no. 7, pp. 750-773, 2003.

[21] E. Kelly, P. J. O'Connell, and E. Smyth, "The economic returns to field of study and competencies among higher education graduates in Ireland," *Economics of Education Review,* vol. 29, no. 4, pp. 650-657, 2010.

[22] M. J. Freire Seoane, M. Teijeiro Álvarez, and C. Pais Montes, "La adecuación entre las competencias adquiridas por los graduados y las requeridas por los empresarios," 2013.

[23] P. Kellermann, "Acquired and Required Competencies Of Graduates," in *Careers of university graduates: Views and experiences in comparative perspectives*, vol. 17, U. Teichler, Ed. Dordrecht: Springer Science & Business Media, 2007, pp. 115-131.

[24] P. Kellermann and G. Sagmeister, "Higher education and graduate employment in Austria," *European Journal of education,* vol. 35, no. 2, pp. 157-164, 2000.

[25] H. Schomburg and U. Teichler, *Higher education and graduate employment in Europe: results from graduates surveys from twelve countries*. Springer Science & Business Media, 2007.

[26] U. Teichler, "Research on the relationships between higher education and the world of work: Past achievements, problems and new challenges," *Higher Education,* vol. 38, no. 2, pp. 169-190, 1999.

[27] U. Teichler, "Graduate employment and work in selected European countries," *European Journal of Education,* vol. 35, no. 2, pp. 141-156, 2000.

[28] U. Teichler, "New perspectives of the relationships between higher education and employment," *Tertiary Education & Management,* vol. 6, no. 2, pp. 79-92, 2000.

[29] W. McKinney, *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. " O'Reilly Media, Inc.", 2012.

[30] W. McKinney, "Data structures for statistical computing in python," in *Proceedings of the 9th Python in Science Conference*, 2010, vol. 445, pp. 51-56: SciPy Austin, TX.

[31] H. Wickham, "Tidy data," *Journal of Statistical Software,* vol. 59, no. 10, pp. 1-23, 2014.

[32] J. Cruz-Benito. (2017). *Jupyter notebook developed to support the research presented in the paper "Proposing a machine learning approach to analyze and predict employment and its factors"* Available: https://github.com/juan-cb/paper-ieeeAccess-2017

[33] S. Raschka, *Python machine learning*. Packt Publishing Ltd, 2015.

[34] L. Breiman, "Random Forests," *Machine Learning,* journal article vol. 45, no. 1, pp. 5-32, October 01 2001.

[35] W. McKinney. (2017). *Pandas, Python Data Analysis Library. 2017*. Available: http://pandas.pydata.org/

[36] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research,* vol. 12, no. Oct, pp. 2825-2830, 2011.

[37] M. Ragan-Kelley *et al.*, "The Jupyter/IPython architecture: a unified view of computational research, from interactive exploration to communication and publication," in *AGU Fall Meeting Abstracts*, 2014.

[38] F. Perez and B. E. Granger, "Project Jupyter: Computational narratives as the engine of collaborative data science," Technical Report. Technical report, Project Jupyter2015.

[39] T. Kluyver *et al.*, "Jupyter Notebooks-a publishing format for reproducible computational workflows," in *ELPUB*, 2016, pp. 87-90.

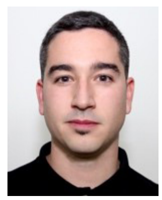
[40] Scikit-learn. (2017, 4/09/2017). *API Reference - scikit-learn documentation: Metrics*. Available: http://scikit-learn.org/stable/modules/classes.html - module-sklearn.metrics

[41] J. Cruz-Benito, A. Vázquez-Ingelmo, J. C. Sánchez-Prieto, R. Therón, F. J. García-Peñalvo, and M. Martín-González, "Enabling Adaptability in Web Forms Based on User Characteristics Detection Through A/B Testing and Machine Learning," *IEEE Access,* vol. 5, 2017.

[42] A. Zollanvari, R. C. Kizilirmak, Y. H. Kho, and D. Hernández-Torrano, "Predicting Students' GPA and Developing Intervention Strategies Based on Self-Regulatory Learning Behaviors," *IEEE Access,* 2017.

Francisco J. García-Peñalvo

Francisco J. García-Peñalvo received the degrees in computing from the University of Salamanca and the University of Valladolid, and a Ph.D. from the University of Salamanca. He is currently the Head of the Research Group Interaction and eLearning. He has led and participated in over 50 research and innovation projects. He was the Vice Chancellor of Innovation with the University of Salamanca from 2007 to 2009. He has authored over 300 articles in international journals and conferences. His main research interests focus on eLearning, computers and education, adaptive systems, Web engineering, semantic Web, and software reuse. He was a Guest Editor of several special issues of international journals, such as Online Information Review, Computers in Human Behavior, and Interactive Learning Environments. He is also a member of the Program Committee of several international conferences and a Reviewer of several international journals. He is currently the Editor-in- Chief of the International Journal of Information Technology Research and the Education in the Knowledge Society Journal. He is also the Coordinator of the multidisciplinary Ph.D. Programme on Education in the Knowledge Society.



Juan Cruz-Benito

Juan Cruz-Benito received an M.Sc. degree in intelligent systems from the University of Salamanca, Spain, in 2013, where he is currently pursuing a Ph.D. in computer sciences. He is one of the youngest members of the Research Group Interaction and eLearning, where he specializes in software solutions based on technology ecosystems and open source software. He works in Human-Computer Interaction, educational virtual worlds and technologies for educational purposes, disciplines that he has developed in many innovation and research projects. He has participated in many European and national R&D projects, such as TRAILER, VALS, USALSIM Virtual Campus, and the Spanish Observatory for University Employability and Employment (OEEU), where he participated as a software engineer, researcher, and developer.

Martín Martín-González

Martín Martín-González is a researcher at the UNESCO Chair in University Management and Policy of the Technical University of Madrid (UPM). Prior to his work at the UPM he worked as a researcher in the Faculty of Economics in the Autonomous University of Madrid (UAM). He holds a PhD in Economics and a master's degree in Economic Development and Public Policy from the UAM. He majored in economics at the University of La Laguna (ULL). His areas of research are the Economics of Education, the Economics of Higher Education, Employability, the Evaluation of Educational Policies, Public Economics, Applied Economics, Economic Development, Higher Education and Vocational Training.

Andrea Vázquez-Ingelmo

Andrea Vázquez-Ingelmo was born in Salamanca, Castilla y León, Spain in 1994. She received a bachelor's degree in Computer Engineering from the University of Salamanca, Salamanca, in 2016, and she is currently completing her master's degree in computer engineering from the same university. She is a member of the Research Group of Interaction and eLearning (GRIAL). Since 2016, she has been part of the national project "Spanish Observatory for University Employability and Employment" as a developer and researcher.

José Carlos Sánchez-Prieto

José Carlos Sánchez-Prieto received a bachelor's degree in Pedagogy and a master's degree in ICT applied in education from the University of Salamanca (Spain). Currently he is a Researcher in Training at the Faculty of Education of said university, where he conducts his PhD research within the Programme on Education in the Knowledge Society. His area of research is the assessment of attitudes among in-service and pre-service teachers.

Roberto Therón

Roberto Therón studied Computer Science at the University of Salamanca (Diploma) and the University of La Coruña (BA). After joining the Research Group Robotics at the University of Salamanca, he presented his thesis work, "Parallel calculation configuration space for redundant robots," receiving the Extraordinary Doctoral Award. He subsequently obtained a Bachelor in Communication Studies (University of Salamanca) and a Bachelor in Humanities (University of Salamanca). At the same university, he continues to carry out his research work as the manager of the VisUsal group (within the Recognized Research Group GRIAL), which focuses on the combination of approaches from computer science, statistics, Graphic Design and Information Visualization to obtain an adequate understanding of complex data sets. In recent years, he has been dedicated to developing advanced visualization tools for multidimensional data, such as genetics or paleo-climate data. In the field of Visual Analytics, he develops productive collaborations with groups and institutions internationally recognized as the Laboratory of Climate Sciences and the Environment (France) or the Austrian Academy of Sciences (Austria). He is the author of over 100 articles in international journals and conferences.

6.13 Appendix M. A Deep-Learning-based proposal to aid users in Quantum Computing programming

J. Cruz-Benito, I. Faro, F. Martín-Fernández, R. Therón and F. J. García-Peñalvo, "A Deep-Learning-based proposal to aid users in Quantum Computing programming," presented in HCI International 2018, Las Vegas, USA, 2018. doi: 10.1007/978-3-319-91152-6_32.

A Deep-Learning-based proposal to aid users in Quantum Computing programming

Juan Cruz-Benito[1[0000-0003-2045-8329]], Ismael Faro[1], Francisco Martín-Fernández[1], Roberto Therón[2[0000-0001-6739-8875]], Francisco J. García-Peñalvo[2[0000-0001-9987-5584]]

[1] IBM Research. T.J. Watson Research Center, Yorktown Heights, NY, USA
[2] GRIAL Research Group. Department of Computer Science, University of Salamanca. Salamanca, Spain
{juan.cruz, ismael.faro1}@ibm.com, pacom@us.ibm.com, {theron, fgarcia}@usal.es

Abstract. New languages like Open QASM and SDKs like QISKit open new horizons for the research and development in the new paradigm of quantum computing. Despite that, they present an evident learning curve that could be hard for regular developers and newcomers in the field of quantum computing. On the other hand, currently there are many ways to build intelligent systems that can learn from humans and processes to build a knowledge corpus and provide a different kind of help to humans in tasks like aiding in decision making processes, recommending multimedia resources, building conversational agents, etc. In this paper we describe a work-in-progress project developed by the IBM Q team that implements an intelligent system based on a deep learning approach that learns how people code using the Open QASM language to later offer help and guidance to the coders by recommending different code sequences, logical steps or even small pieces of code. During the paper, we describe our current approach and first results. They include the use of *seq2seq* neural networks that effectively learn quantum-code sequences, and which will be tested in real context in the near future to improve the user experience in IBM Q Experience products.

Keywords: Deep Learning, Artificial Intelligence, Quantum Computing, Programming, Open QASM, QISKit.

1 Introduction

Quantum computing programming is not currently an easy task. New languages like Open QASM [1] and SDKs like QISKit [2, 3] open new horizons for the research and development in the new paradigm of quantum computing [4, 5]. Despite that, they present a non-easy learning curve for regular developers and newcomers in the field of quantum computing. On the other hand, there are nowadays many ways to build intelligent systems that can learn from humans and build a knowledge corpus on their own. These knowledge corpuses could be used to provide a different kind of help to humans in tasks like aiding in decision making processes, recommending multimedia resources, building conversational agents, etc.

Related to the aforementioned intelligent systems, we find in the literature and new media buzzwords like Artificial Intelligence (AI) [6], Machine Learning (ML) [7, 8], Deep Learning (DL) [9, 10], etc. Many modern applications include these kind of concepts and keywords to look trendy; others involve them to deal with problems that are difficult to solve in other traditional ways. Apart of the trendiness of the terms, it is a fact that these research fields are increasingly present: many enterprises are spending a lot of effort and money to be AI-driven; including AI in applications, decision systems, etc. In this sense and related to the concept of aid provided by AI or intelligent systems, we present our core-concept of User Experience (UX). The UX is commonly defined as “a person’s perceptions and responses that result from the use or anticipated use of a product, system or service” [11]. Related to the ISO definition, UX includes all the users’ emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviors and accomplishments that occur during, before and after the usage. According to this ISO, there are three factors that influence user experience: system, user and the context of use. Considering these three factors, and related to the new wave related to AI, we can think that it is possible to interfere on the system or the context of use [12] by applying AI to existing systems. Thinking in the AI application, we can help to improve the UX, since the AI could learn from users’ and previous usage to change or adapt the system or specific features to the current user’s needs, desires and behaviors.

Many times, the UX is merely considered in the context of visual applications (visual UIs) or related to common products designed for regular users. In the case of our work-in-progress research, we try to improve the user experience in the context of programming under the quantum computing paradigm. Learning how people code using the products developed by IBM Research [13], we think is possible to distill knowledge to later use it in guiding and helping other quantum-programmers in common tasks related to the code. In our experience, currently there are many common issues on coding quantum programs (definition of qubits to use, measurement operations, etc.) based on some common rules and patterns, that could represent general (and simple) cases where the users could initially be helped. This help could positively affect the users by fulfilling their desires and expectations to achieve success with their code and experiments or on creating more positive experiences through being helped by an intelligent system that could provide real-time feedback on their code [14, 15].

Gathering all these ideas and core concepts, in this paper we describe a work-in-progress (WIP) project developed by the IBM Q team that implements an intelligent system based on a deep learning approach that learns how people code using the Open QASM language to offer help by recommending different code sequences, logical steps or even small pieces of code. To present this work, section two introduces the core concepts of the *seq2seq* approach and the importance of the natural language processing (NLP) in our context. The third section presents our proposal, describing how our WIP project is developed in terms of technology, details on datasets, etc. as well as our first results achieved. The fourth and fifth sections depict the future work to be done and a brief conclusion on our proposal.

2 *Seq2seq* and the importance of NLP approaches

What is source code? In fact, and with no intention of providing a deep definition, the code is a human-readable set of words or alphanumerical characters (instructions) previously defined that could be accompanied by other characters like punctuation, etc., and follows a logical structure and some kind of grammar [16, 17]. Following this consideration (and the *programming languages* idea) we find that the foundations of coding are not so far from those that define the human languages.

In the natural language processing area (NLP) many researchers work using artificial intelligence to analyze human language and design conversational systems, to summarize automatically texts, to learn and replicate communicative styles, etc. In this sense we are using concepts from NLP to teach neural networks how to code using quantum computing languages and libraries (mainly using Open QASM [1]) like the human programmers. That is, we are feeding recurrent neural networks (RNN), as we will explain below, the code entered by programmers when they use the IBM Q Experience [18] to enable them to learn how the code is composed in the context of quantum programming.

Regarding the approach we are following under the NLP umbrella, it is the sequence-to-sequence (*seq2seq*) neural network model [19-21]. This model consists of two RNN's: "one RNN encodes a sequence of symbols into a fixed-length vector representation, and the other decodes the representation into another sequence of symbols. The encoder and decoder of the proposed model are jointly trained to maximize the conditional probability of a target sequence given a source sequence" [20]. That is, using this method is possible to train a system that produces sequences of symbols using an input sequence of symbols (Fig. 1).

This approach has been used recently with success in the tasks of performing translations between different languages [22, 23] (Fig. 1), since its ability of learn semantically and syntactically meaningful representation of linguistic phrases [20].

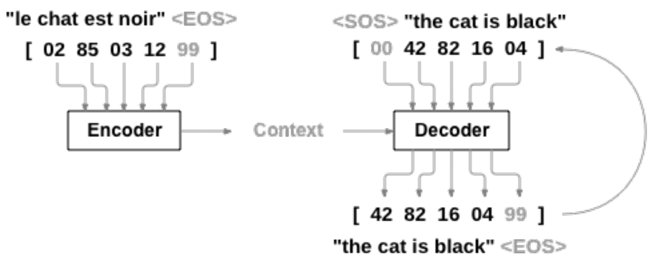


Fig. 1. Overview of a sequence to sequence network using different networks as encoders and decoders. Image taken from [24]

In our case, the *seq2seq* neural network is not employed to translate languages, but it is used similarly to translate some sequences to other ones. In our approach, the input sequence will be the one typed by the user in the tools for developing quantum code, and the target sequence will be next logical sentence(s) proposed by the neural networks. This will be explained further in the following section.

3 Proposal and first results

Based on previous experiences by the authors [25] and other researchers [14, 26], the main idea of this paper is that intelligent systems could enhance the user experience. Furthermore, they could enhance the experience of developers while they code new programs in challenging environments like ours. As previously stated, our goal is to provide real-time feedback to IBM Q users by proposing code to them in the different quantum programming environments developed by IBM Research and IBM Q Experience team [18]. In order to provide the recommendations to the users, we are developing a neural network based on a *seq2seq* approach that learns the sequences from the code composed by the users to develop quantum programs in both ways: learning what kind of sentences are used (and relevant for quantum computing) and also the logical sequences they follow to build a quantum program.

To pursue this, there are some important factors to keep in mind: how to develop the *seq2seq* network and its different utilities and how to train it using the data available in our platforms.

First of all, regarding to the technological issues, we use PyTorch [27] as our deep learning framework, supported by the *PyTorch-seq2seq* [28] library from IBM, which is used to build our *seq2seq* network. We selected PyTorch due its features related to the creation of dynamic neural networks and its performance on building deep learning models using GPUs. Also, we use the *seq2seq* library for PyTorch developed by our colleagues at IBM to work from a tested approach in terms of the *seq2seq* models and to avoid starting from scratch for our WIP. This library encloses different functionalities related to the RNN that encodes the input sequences and the decoder RNN that produces the output (target) sequences. Apart of the functionalities it has, we have developed a variation (available in https://github.com/IBM/pytorch-seq2seq/pull/116) on the prediction method used to produce the target sequences. This variation on the prediction method not only produces a unique prediction based on the most probable output sequence (as in the original implementation) but also includes a beam search strategy [19, 29] to produce several possible outputs (target sequences) with variations given an input sequence. In fact, with this change we are replacing the default RNN decoder to use a TopKDecoder RNN.

Secondly, we developed datasets to train our models in the task of predicting the proper code sequence that would be result in adding the next logical sentence to the given code sequence. We built our datasets using the following rule *n -> n+1*: given a code sequence 'n', the predicted sequence would be 'n+1', where 'n' is the set of instructions that compose the input sequence, and the ‘+1’ is the following instruction that could be used after the last instruction in the set ‘n’.

The dataset development consisted of two main phases: (1) designing the code sequences to represent inputs and outputs (given and target sequences) and (2) building simpler representations of the code sequences to facilitate the training. On the designing of the code sequences, we augmented the original dataset of quantum programs sent to the IBM Q backends by dividing the quantum program into all the possible parts that followed the rule *n -> n+1* (input -> output sequence). To reduce the complexity of the training phase, we replaced each unique instruction using a unique key, thereby mapping the instructions used in the code developed by the users to simpler representations. This unique key (composed by 1-3 alphanumerical characters typically), will be used in the datasets to train the encoder and decoder RNN. To build the mapping, we parsed all the defined Open QASM statements and their possible arguments. The list of these Open QASM statements is available in the Fig. 2.

Statement	Description	Example
OPENQASM 2.0;	Denotes a file in Open QASM format	OPENQASM 2.0;
qreg name[size];	Declare a named register of qubits	qreg q[5];
creg name[size];	Declare a named register of bits	creg c[5];
include “filename”;	Open and parse another source file	include “qelib1.inc”;
gate name(params) qargs	Declare a unitary gate	(see text)
opaque name(params) qargs;	Declare an opaque gate	(see text)
// comment text	Comment a line of text	// oops!
U(theta,phi,lambda) qubit|qreg;	Apply built-in single qubit gate(s)	U(pi/2,2*pi/3,0) q[0];
CX qubit|qreg,qubit|qreg;	Apply built-in CNOT gate(s)	CX q[0],q[1];
measure qubit|qreg -> bit|creg;	Make measurement(s) in `z` basis	measure q -> c;
reset qubit|qreg;	Prepare qubit(s) in `|0\rangle`	reset q[0];
gatename(params) qargs;	Apply a user-defined unitary gate	crz(pi/2) q[1],q[0];
if(creg==int) qop;	Conditionally apply quantum operation	if(c==5) CX q[0],q[1];
barrier qargs;	Prevent transformations across this source line	barrier q[0],q[1];

Fig. 2. Open QASM language statements (version 2.0). Figure from [1, 30]

To provide an example of how we are building the dataset main (later separated into train/test/dev datasets) using our *n -> n+1* rule and the statements mapping, we will use an Open QASM implementation of *Deutsch–Jozsa algorithm* [31] using two qubits (as it appears in [30]).

```
OPENQASM 2.0;
include "qelib1.inc";
qreg q[5];
creg c[5];
x q[4];
h q[3];
h q[4];
cx q[3],q[4];
h q[3];
measure q[3] -> c[3];
```

After mapping the code to the simpler keys, we would get the following code (we have removed the *line feed* between the different lines, making the code a single line):

```
O; I; Q5; C5; X4; H3; H4; CX34; H3; M33;
```

Later, we augment for the dataset the mapped code in different lines, where each line follows the *n -> n+1* rule, and a line below another represents *n = n-1*. So, using the mapped code, we will get the following logical code sequences (Table 1).

Table 1. All the possible logical code sequences for the mapped *Deutsch–Jozsa algorithm* following our augmentation rules

Input sequence (given one)	**Output sequence (target)**
O; I; Q5; C5; X4; H3; H4; CX34; H3;	O; I; Q5; C5; X4; H3; H4; CX34; H3; M33;
O; I; Q5; C5; X4; H3; H4; CX34;	O; I; Q5; C5; X4; H3; H4; CX34; H3;
O; I; Q5; C5; X4; H3; H4;	O; I; Q5; C5; X4; H3; H4; CX34;
O; I; Q5; C5; X4; H3;	O; I; Q5; C5; X4; H3; H4;
O; I; Q5; C5; X4;	O; I; Q5; C5; X4; H3;
O; I; Q5; C5;	O; I; Q5; C5; X4;
O; I; Q5;	O; I; Q5; C5;
O; I;	O; I; Q5;
O;	O; I;

Using this technique, we think we will be able to teach the *seq2seq* network what are the logical steps in terms of Open QASM statements (and their arguments) used to code a quantum program, in a similar mode to the *seq2seq* training to translate between different languages.

To test this approach, we built some small datasets (typically including several thousands of code sequences) to train our *seq2seq* networks and validate our main idea. In the case of these first tests, we used a mapping algorithm that discards the arguments that accompanied the statements in the original code. This could be useful to validate if the *seq2seq* networks are learning the code grammar properly (which could be easier to discover considering the small size of the initial datasets tested). For example, in a typical quantum program, after defining a register of qubits we defined a register of classical bits. In our tests, we do not mind if the number of qubits is the same as the number of bits or other details. We just want to validate if the neural network has learned that a declaration of a register of classical bits should follow a declaration of a register of qubits. The initial results with the training, using the simpler mapping, show that our *seq2seq* network achieves a prediction accuracy of 88-92% by comparing the training and test datasets (using the default decoder predictor of just one statement more in the target sequence per each sequence input). In our case, we also introduced a quantitative testing of the results provided by the dataset. In this qualitative testing we tried to verify if the results provided by the default predictor and the predictor based on beam search are comprehensive and logical in the context of quantum programming. This qualitative testing is currently ongoing, and it involves some IBM Q team members with expertise in quantum coding who will provide their impressions on the results achieved.

Also, as part of our initial results, we have designed the two main approaches to deploy our intelligent system in real contexts to integrate the assistant in the IBM Q products. These two approaches are the following:

1. Build an API to serve the results of the predictions and submit new codes to continue training our *seq2seq* network. In this case, we follow a *deep learning as a service* approach. The neural network and the different features are available in the cloud through using a REST API (currently implemented using Flask in our case). Using this solution, the different products that will include the intelligent code recommender solution only need to submit the contents of the quantum program being written by the user to obtain what would be the next statement to use.
2. Embed the trained model within the products itself. Considering that the training model is saved into a *.pt* file of about 3MB, we can embed the file within a website or a program and use it by employing ONNX [32] or introducing our code related to the predictions.

As stated, of these two approaches, we have built only the deployment with the REST API. It will be used in the full validation of the system. In the other case of the embedded code, implementing it will depend on the final success of our full tests.

4 Next steps and future work

The next steps that we will take will be the following:

1. Finalize the internal validation of the simpler version of the Open QASM recommender. This will help us to understand whether the predictions fulfill our expectations or if we need to redefine our intelligent system (tuning the seq2seq network, changing the neural network used, etc.).
2. Build a full dataset to train the neural network. IBM Q has over 2 million records of code executions. We plan to build an augmented dataset using the *n -> n+1* rule and include all the code introduced by users to train our future neural networks.
3. Validate internally at IBM Q the final results of these full trainings. At this phase, we will validate the results raised by the neural networks using the

simple mapping as well as using a complex mapping (which maps all the statements and their arguments).

4. If all those validations and trainings are successful, we plan to deploy our code recommender using the API REST to be used by the other products and explore the integration of our code into other IBM Q products.
5. Using the deployed version in real products, we will measure different metrics of real user experiences to assess the effect of the intelligent system. In this case, we will measure (prospectively) the ratio between the code offered by the system and those statements employed finally by the users, the time they use to complete their programs using the helper, and the users' opinion through questionnaires, focus groups or other assessment tools typically used in the Human-Computer Interaction research area.
6. Finally, the intelligent system should be re-trained regularly to adapt its knowledge to the new code produced by users. In a novel context like quantum programming, it is possible that the number of users coding as well as the complexity of the code produced will grow. The system will need to be ready to fulfill the expectations of different users with different level of coding skills in quantum computing.

5 Conclusions

This paper describes a work-in-progress project that implements an intelligent system based on a deep learning approach to learn how people code in Open QASM language. This knowledge acquired by the neural networks will be used to offer help and guidance to the programmers by recommending different code sequences, logical steps or even small pieces of code. We intend that this help and guidance during the programming will improve the user experience (UX) of quantum-programmers within the IBM Q products. During the paper we have described our ideas, our current implementation and the different initial results achieved. Among these initial results, we highlight that we have developed *seq2seq* neural networks that are learning quantum code sequences following custom datasets built on our own specification designed to predict possible code sequences given the code produced by users previously. We also provide real examples of how the datasets will be produced and the accuracy achieved by our system in simple test cases. Finally, we described our next steps and the future work that we will try to tackle in the near future.

6 Acknowledgement

We thank to the ACQX (AI Challenges & Quantum Experience Team at IBM Research) and to the GRIAL research group team for the useful discussions and the feedback received.

References

1. Cross, A.W., Bishop, L.S., Smolin, J.A., Gambetta, J.M.: Open quantum assembly language. arXiv preprint arXiv:1707.03429 (2017)
2. Cross, A.: The IBM Q experience and QISKit open-source quantum computing software. Bulletin of the American Physical Society (2018)
3. IBM Research, QISKit. Quantum Information Software Kit, https://www.qiskit.org/, last accessed 18/2/2018
4. Kandala, A., Mezzacapo, A., Temme, K., Takita, M., Brink, M., Chow, J.M., Gambetta, J.M.: Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. Nature 549, 242 (2017)
5. Ristè, D., da Silva, M.P., Ryan, C.A., Cross, A.W., Córcoles, A.D., Smolin, J.A., Gambetta, J.M., Chow, J.M., Johnson, B.R.: Demonstration of quantum advantage in machine learning. npj Quantum Information 3, 16 (2017)
6. Nilsson, N.J.: Principles of artificial intelligence. Morgan Kaufmann (2014)
7. Raschka, S.: Python machine learning. Packt Publishing Ltd (2015)
8. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V.: Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 12, 2825-2830 (2011)
9. Goodfellow, I., Bengio, Y., Courville, A.: Deep learning. MIT press (2016)
10. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature 521, 436-444 (2015)
11. ISO/DIS: 9241-210: 2010. Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems. International Standardization Organization (ISO). Switzerland (2009)
12. ISO/DIS: Draft BS ENISO 9241-220 Ergonomics of human-computer interaction. Part 220: Processes for enabling, executing and assessing human-centred design within organizations. (2016)
13. Castelvecchi, D.: Quantum cloud goes commercial. Nature 543, 159-159 (2017)
14. Lee, S., Choi, J.: Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity. International Journal of Human-Computer Studies 103, 95-105 (2017)
15. Shneiderman, S.B., Plaisant, C.: Designing the user interface 4 th edition. Pearson Addison Wesley, USA (2005)
16. The Linux Information Project, Source Code Definition, http://www.linfo.org/source_code.html, last accessed 20/2/2018
17. Harman, M.: Why source code analysis and manipulation will always be important. In: Source Code Analysis and Manipulation (SCAM), 2010 10th IEEE Working Conference on, pp. 7-19. IEEE, (2010)
18. IBM Research, The IBM Quantum Experience, https://www.research.ibm.com/ibm-q/, last accessed 18/2/2018
19. Sutskever, I., Vinyals, O., Le, Q.V.: Sequence to sequence learning with neural networks. In: Advances in neural information processing systems, pp. 3104-3112. (2014)
20. Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y.: Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078 (2014)

21. Vinyals, O., Le, Q.: A neural conversational model. arXiv preprint arXiv:1506.05869 (2015)

22. Bahdanau, D., Cho, K., Bengio, Y.: Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473 (2014)

23. Neubig, G.: Neural machine translation and sequence-to-sequence models: A tutorial. arXiv preprint arXiv:1703.01619 (2017)

24. Robertson, S., Translation with a Sequence to Sequence Network and Attention, http://pytorch.org/tutorials/intermediate/seq2seq_translation_tutorial.html, last accessed 20/2/2018

25. Cruz-Benito, J., Vázquez-Ingelmo, A., Sánchez-Prieto, J.C., Therón, R., García-Peñalvo, F.J., Martín-González, M.: Enabling Adaptability in Web Forms Based on User Characteristics Detection Through A/B Testing and Machine Learning. IEEE Access 6, 2251-2265 (2018)

26. Loup-Escande, E., Frenoy, R., Poplimont, G., Thouvenin, I., Gapenne, O., Megalakaki, O.: Contributions of mixed reality in a calligraphy learning task: Effects of supplementary visual feedback and expertise on cognitive load, user experience and gestural performance. Computers in Human Behavior 75, 42-49 (2017)

27. Paszke, A., Gross, S., Chintala, S., Chanan, G., PyTorch. Tensors and Dynamic neural networks in Python with strong GPU acceleration, https://github.com/pytorch/pytorch, last accessed 20/2/2018

28. IBM Research, pytorch-seq2seq, https://github.com/IBM/pytorch-seq2seq, last accessed 20/2/2018

29. Wiseman, S., Rush, A.M.: Sequence-to-sequence learning as beam-search optimization. arXiv preprint arXiv:1606.02960 (2016)

30. IBM Research, Open QASM. Gate and operation specification for quantum circuits. , https://github.com/QISKit/openqasm, last accessed 20/2/2018

31. Deutsch, D., Jozsa, R.: Rapid solution of problems by quantum computation. In: Proc. R. Soc. Lond. A, pp. 553-558. The Royal Society, (1992)

32. Open Neural Network Exchange, ONNX Github repository, https://github.com/onnx/onnx, last accessed 20/2/2018