# 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.
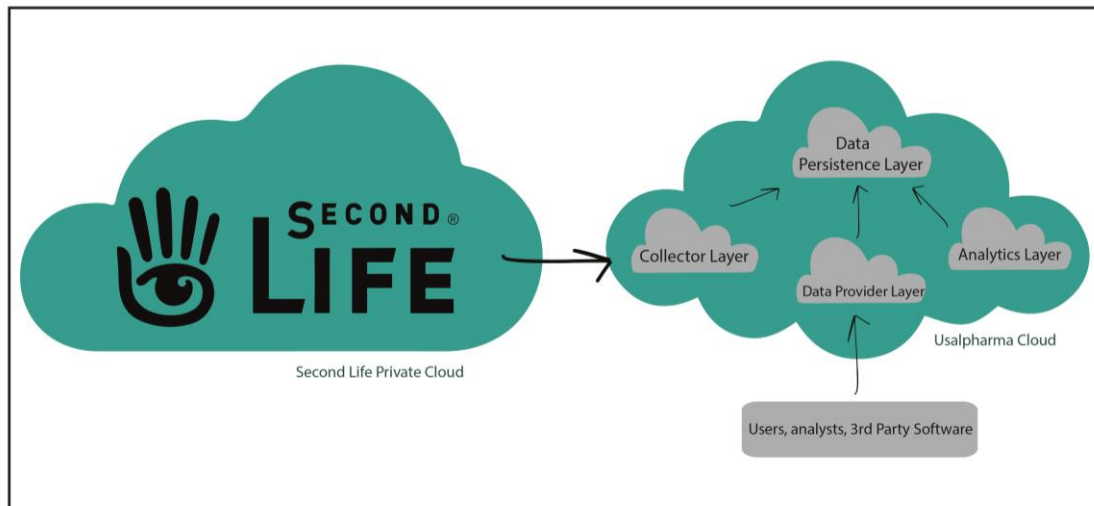


Figure 1: Cloud Architecture divided into functionality layers [7]

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.

## Checkpoints

### username

Total checkpoints achieved 1 / 32
3.1% checkpoints revised

## Rooms:

0/8 Rooms accesed, 0% of total rooms

Rooms not accesed:
- Puerta Muestroteca
- Puerta Despacho
- Puerta Reactivos
- Puerta Residuos
- Salida Residuos
- Puerta Sas
- Acceso Lab

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 destined 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.