



VNiVERSiDAD D SALAMANCA

Privacidad y gestión de la identidad en procesos de analítica de aprendizaje

TESIS DOCTORAL

Programa de Doctorado Formación en la Sociedad del Conocimiento de la Universidad de Salamanca

Doctorando:

Daniel Amo Filvà

Directores:

Dr. Francisco José García-Peñalvo

Dr. Marc Alier

Dr. David Fonseca

Diciembre, 2019

Resumen

Las analíticas en el contexto educativo (*Academic Analytics / Learning Analytics*) implican un procedimiento de explotación de datos para la mejora del proceso de enseñanza/aprendizaje. Este procedimiento consiste en recolectar, analizar y crear visualizaciones de los datos de los procesos educativos y/o de los estudiantes. Los datos personales, registros de actividad y metadatos de estudiantes y profesores se almacenan, comparten, transforman y utilizan a discreción por las instituciones educativas y servicios de terceros.

La privacidad, confidencialidad y seguridad de los datos personales de estudiantes quedan expuestas a diario cuando no hay un control o gestión adecuado. El tratamiento analítico de los datos puede ir en contra de los intereses o voluntad de los estudiantes. Esta situación es más delicada cuando se involucran menores de edad. El uso de las analíticas en educación está generando un creciente entorno de desconfianza en cuanto el tratamiento de datos de las personas involucradas.

Los procesos educativos donde intervienen procedimientos de analíticas educativas presentan un doble problema. Por un lado, la fragilidad de los datos debida a la baja protección de la privacidad, la confidencialidad y la seguridad de datos en los almacenes digitales. Por otro lado, la falta de madurez en los procedimientos y soluciones de protección de los datos personales y de la identidad de los estudiantes. El problema es grave y afecta tanto al uso como a la transferencia y custodia de datos generados por el estudiante.

Los términos de uso y políticas de privacidad imponen unas condiciones que estudiantes y profesores deben validar para usar el servicio. No obstante, durante el uso del servicio no suele quedar claro ni para los estudiantes, ni para los profesores, ni para las instituciones dónde están los límites de uso, acceso, gestión o tratamiento de datos.

Las leyes de protección de datos otorgan una serie de derechos a estudiantes y profesores, contemplando incluso situaciones excepcionales y de índole personal. Estos derechos ofrecen un margen de libertad en la configuración de sus perfiles, que las plataformas de aprendizaje deben considerar como parte de su diseño y activadas por

defecto. Las plataformas educativas no disponen de las funcionalidades para ejercer todos los derechos.

La presente tesis investiga de forma analítica el estado de la cuestión. El trabajo realizado en la investigación identifica, diseña y evalúa soluciones que resuelven total o parcialmente de la problemática descrita. En primer lugar, se exploran las posibilidades de la tecnología emergente *blockchain*. En segundo lugar, se evalúan soluciones a nivel de almacén de datos en los entornos virtuales de aprendizaje. En ambas perspectivas se aborda una parte experimental centrada en el desarrollo de prototipos funcionales para:

- Que exista una adecuada protección, confidencialidad y seguridad de los datos educativos almacenados.
- Que las plataformas educativas estén tecnológicamente preparadas para asumir el ejercicio de los derechos del estudiante.
- Y, finalmente, que se pueda transferir a roles educativos la necesidad de proteger a los estudiantes y profesores haciendo un uso correcto de las herramientas digitales del aula.

Los resultados de la investigación reflejan que el problema es complejo y múltiple. Se demuestra que *blockchain* no puede aportar soluciones a la privacidad y seguridad a nivel del almacén de datos. Desde la perspectiva de los entornos virtuales de aprendizaje, se aportan propuestas para el avance de la ciencia y se implementan soluciones funcionales a cuestiones concretas del problema. La investigación aporta una evolución clara al estado de la cuestión y nuevas líneas de trabajo en las que abordar distintas cuestiones de la problemática a futuros.

Palabras clave: Blockchain, Smart Contracts, Learning Analytics, Entornos Virtuales de Aprendizaje, Confidencialidad de Datos Personales, Gestión de la Identidad Digital, Reglamento General de Protección de Datos, Privacidad, Seguridad.

Abstract

Analytics in the educational context (*Academic Analytics / Learning Analytics*) involves an analytical procedure that aims to potentially improve the teaching/learning process by collecting, analyzing, and creating visualizations of data from educational processes and/or students. During the process, students' and teachers' data and metadata are stored, shared, transformed, and used at the discretion of educational institutions and third-party services, a treatment that may be against the students' interests or will. Without proper control or management, the privacy, confidentiality, and security of students' data are exposed on a daily basis, which is a particularly delicate issue when it comes to underage students. Therefore, the use of analytics in education generates an environment of mistrust regarding the processing of data of the persons involved.

As a result of the above, there is a serious problem of fragility, or lack of maturity in the state of the question, concerning the protection of personal data and student identity in educational processes where educational analytical procedures intervene. As the results of this research will demonstrate, there is low protection of privacy, confidentiality, and data security in digital repositories, affecting not only the use, but also the transfer and custody of student-generated data.

It should be noted that the terms of use and privacy, policies impose conditions that students and teachers must validate to use the service. However, during the use of the service, it is often not clear to students, teachers, or institutions where the limits of data use, access, management, or processing lie. Data protection laws grant several rights that can be exercised by students and teachers, even in exceptional and personal situations. These rights offer a margin of autonomy in the configuration of profiles, which learning platforms must consider as functionalities by design and by default. In the research carried out, it is easy to find examples of the shortcomings in the current platforms.

This thesis focuses on an analytical investigation of the state of the matter, with the main objective of proposing solutions to resolve this fragile situation in the processing

of personal data from an information and technology policy perspective, based on the following premises:

- That the levels of confidentiality and security of the stored educational data are adequate.
- That educational platforms are technologically prepared to incorporate the exercise of student rights.
- And, finally, to foment the need among educational institutions to protect students and teachers and make proper use of digital classroom tools.

The research work carried out attempts to identify, design, and evaluate solutions to the problems described from two different perspectives. In the first place, it explores the solving possibilities of emerging technology *blockchain*. Secondly, it looks for and evaluates existing solutions in the technological environment of data warehouses in virtual learning environments. Both perspectives address an experimental part focused on the development of functional prototypes that seek to solve all or part of the initial problems identified.

The results of the investigation reflect that the problem is complex and multiple, demonstrating that *blockchain* cannot provide privacy and security solutions at the level of the data warehouse. At the same time, and from the perspective of virtual learning environments, the results provide innovative evidence that helps solve specific problem issues and advance science. This research contributes to the state of the question by providing new insights and lines of work which address different both current and future concerns of analytics in the educational context.

Keywords: Blockchain, Smart Contracts, Learning Analytics, Virtual Learning Environments, Personal Data Confidentiality, Digital Identity Management, General Data Protection Regulation, Privacy, Security.

Agradecimientos

Escribir mi tesis doctoral ha sido un viaje de pequeños viajes. Una biografía se empieza a escribir cuando se ha acumulado un conjunto de experiencias. Se puede decir lo mismo de una tesis doctoral. Se escribe cuando has vivido una serie de experiencias de investigación, y digo serie en lugar de conjunto ya que es un proceso ordenado, lógico y organizado mientras pasan otras cosas alrededor. Escribirla ha sido como revivir todas las experiencias acaecidas durante el trayecto, desde que empezaron las primeras conversaciones hasta el momento de escribir justamente este apartado. Y de las experiencias se aprende. Se aprende a ser crítico, a dialogar, a escribir, a ser riguroso, a ser organizado... pero sobretodo se aprende a ser paciente, reflexivo y autocrítico. Sin duda alguna, la tesis doctoral es un crecimiento académico, pero también personal.

Como dijo Ortega y Gasset, “yo soy yo y mi circunstancia”. Durante la tesis doctoral he vivido distintas circunstancias, que me han hecho ser el yo cargado de la circunstancia de vivir el final de la tesis doctoral, de escribirla. Mi circunstancia de ahora es el resultado de las pequeñas circunstancias pasadas junto a mi yo. Y estoy orgulloso, así como espero que estén orgullosos a quien ahora les voy a dedicar unas palabras.

Cada una de las circunstancias pasadas las he vivido junto a distintas personas, a las que estoy profundamente agradecido por sus comentarios, reflexiones, ánimos y críticas, o sencillamente por su presencia. Algunas de ellas me han acompañado desde el principio, otras han sido como destellos fugaces que dejan huella capaz de darle un giro brusco a todo lo construido.

Sin lugar a duda estoy inmensamente agradecido con Esther Sànchez, con quien comparto vida y dos hijas como profesoras de la asignatura “ser padre y madre”. Gracias por escucharme literalmente en cualquier momento, sin entender muchas veces, ni yo mismo, lo que le explico, pero que, sin duda, y estoy seguro de que esto sí lo sabe, ha sido un pilar fundamental para llegar al yo que soy ahora.

Mi aprendizaje y evolución de la tesis también lo ha vivido mi familia durante las cenas de Navidad, en los encuentros semanales y cualquier momento compartido, hasta

nadando en los veranos calurosos de Badalona no podía dejar de comentar distintos aspectos. Gracias por no dejarme solo en Telegram.

Por supuesto, mis tres directores lo han sido todo durante el proceso. Estoy muy orgulloso de que sean ellos, además amigos, que la providencia decidió unirlos en una serie de sucesos aparentemente inconexos. Ellos son Marc Alier, Francisco José García-Peñalvo y David Fonseca, ordenados ascendentemente según momento en el que los conocí. Y en este orden voy a agradecerles su dirección y su eterna paciencia.

He tenido la suerte de conocer a Marc Alier y crecer con él desde que era menor de edad. Y lo digo sin tapujos, es mi modelo que seguir. Una persona que se hace llamar (Gran) Ludo, en referencia a ese monstruo lanudo de la película Dentro del Laberinto de 1986, muy parecido a Chewbacca de Star Wars, pero con unos kilos de más y con la capacidad de comunicarse mentalmente con rocas y piedras (Marc, no el Ludo de la película), merece todo mi respeto y atención. Su acompañamiento y capacidad crítica interminable, seguramente debida a mi capacidad de tortura por escribir tan bien, han sido fundamentales para corregir mi corriente multipotencial para centrarme en la tesis y dirigir el rumbo de esta con comentarios precisos y contundentes. Marc es un cirujano del pensamiento.

Francisco José García-Peñalvo es quien representa a la Universidad de Salamanca como institución que sustenta el Programa de Doctorado en el que se ha desarrollado esta tesis doctoral. Le estoy eternamente agradecido por acogerme, por aceptar mi propuesta de tesis y por aportar su profundo conocimiento, bagaje y experiencia en lo académico universitario y campo educativo. Su cátedra, visión, concepción de la realidad educativa, consejos y directrices han sido justamente dichas en el momento adecuado para corregir y redirigir cuando a mí me parecía todo abstracto, confuso y sin final. Como dice la serie televisiva “Better Call Saul”, cuando hay que decidir un tema crucial, entonces “mejor llamar a Fran”.

David Fonseca es mi tercer director de tesis y coordinador de GRETEL, grupo de investigación al que pertenezco dentro de La Salle Barcelona. Conecté con David de alguna manera en el congreso CISTI 2014, donde, aunque no lo conocí personalmente sí me encontré por primera vez con Francisco. Luego, en mi entrada como profesor en La

Salle Barcelona, conocí a David como director de grupo GRETEL. Desde entonces ha sido un gran líder y un buen profesor sobre investigación, del que espero haber sido un buen aprendiz. David es una persona muy próxima, empática, pragmática, resolutiva y con un gran conocimiento del ámbito académico-científico que su cátedra avala. Solo diré que es capaz de dirigir a un grupo de investigación de más de 15 personas, conocer en detalle y resolver las necesidades de cada uno, y solucionar sus responsabilidades académicas sin pestañear. Además, hace unas barbacoas para chuparse los dedos. David es mi gran maestro Yoda todoterreno.

Por supuesto, los amigos y conocidos han sido piedras angulares para construir los cimientos de muchos conceptos elaborados en la tesis. Muchas gracias: a mis mejores amigos Lluís Gea y Marc Estruch, cuyo humanismo profundamente reflexivo del primero y la propiedad tecnófila innata del segundo, han aportado una contraposición capaz de completar muchos vacíos argumentales; a los coautores y profesores que participaron en el libro “Analítica del Aprendizaje: 30 experiencias con datos en el aula”, cuyas experiencias han apuntalado muchas conclusiones extraídas en mi investigación; a Fundación Bias y Grupo MT, liderados por Sofía Temprado y Ignacio Romero, dos personas en mayúsculas que tienen una visión de educación muy especial y a quienes les debo muchas de mis experiencias educativas reflexivas a pie de claustros; a toda la comunidad educativa del mundo real y virtual por hacerme partícipe del contexto educativo, y poder contrastar muchos aspectos aquí presentados; a todas las personas con las que he estado en contacto en La Salle Barcelona Universitat Ramon Llull, Universitat Politècnica de Catalunya, Universitat Oberta de Catalunya y Universidad de Salamanca.

Sé que olvido muchos nombres, que ahora no recuerdo, pero su ayuda no ha sido menos importante.

Índice

I.	Introducción.....	27
I.1.	Acotación del objeto de estudio.....	35
I.2.	Objetivos y preguntas de investigación.....	37
I.2.1.	Objetivos.....	38
I.2.2.	Preguntas de investigación.....	40
I.2.3.	Metodología	40
I.3.	Marco de trabajo	44
I.4.	Organización del documento	45
II.	Marco teórico	46
II.1.	Concepto de privacidad	46
II.1.1.	Privacidad o confidencialidad de datos personales	47
II.1.2.	Seguridad	48
II.2.	Internet insegura y <i>clickstream</i>	49
II.2.1.	Estadios en la evolución de internet	50
II.2.2.	<i>Clickstream</i>	54
II.3.	MOOC	55
II.3.1.	Origen de los MOOC.....	55
II.3.2.	Banco de datos para experimentación: minería de datos	59
II.3.3.	Mejora de rendimientos con <i>Clickstream</i>	61
II.4.	<i>Educational Data Mining</i>	62
II.4.1.	<i>Learning Analytics</i>	63
II.4.2.	<i>Academic Analytics</i>	67
II.4.3.	<i>Social Network Analytics</i>	68

II.4.4.	Miedos y recelos: una cuestión delicada.....	70
II.5.	<i>Learning Analytics</i> y Conocimiento en congresos.....	72
II.5.1.	Congresos <i>Learning Analytics</i> and Knowledge	74
II.6.	Leyes sobre protección de datos personales	78
II.6.1.	Reglamento General de Protección de Datos	80
II.6.2.	Protección y privacidad de datos Unión Europea-Estados Unidos	82
II.6.3.	Conocimiento de las leyes educativas.....	85
II.7.	Blockchain.....	90
II.7.1.	Criptomonedas	92
II.7.2.	Tecnología emergente.....	93
II.7.3.	Educación, <i>Learning Analytics</i> y <i>blockchain</i>	94
II.7.4.	<i>Blockchain</i> y DLT.....	95
II.8.	Revisión sistemática de la literatura.....	97
II.8.1.	Revisión y mapeo sistemáticos.....	99
II.8.2.	Resultados del mapeo sistemático	108
II.8.3.	Resultados de la revisión sistemática.....	115
II.8.4.	Análisis resumen de las soluciones propuestas	134
II.8.5.	Amenazas a la validez de esta revisión de la literatura.....	141
III.	Marco empírico	143
III.1.	Metodología	144
III.2.	Excepciones del RGPD en Moodle.....	144
III.3.	Experiencia de usuario	148
III.3.1.	Primera aproximación	149
III.3.2.	Encuestas a perfiles estudiantes	152
III.3.3.	Entrevistas a perfiles profesores.....	154

III.3.4.	Entrevistas a perfiles administradores técnicos de Moodle	162
III.3.5.	Entrevistas a perfiles Delegados de Protección De Datos.....	166
III.3.6.	Conclusiones de los resultados.....	169
III.4.	Desarrollo de una solución en formato <i>plugin</i>	170
III.5.	Líneas trabajadas en paralelo.....	180
III.5.1.	Seguridad en la tabla de usuarios.....	180
III.5.2.	Seguridad en los registros de interacciones.....	182
III.5.3.	Seguridad en los datos almacenados	183
III.5.4.	Interoperabilidad con <i>blockchain</i>	185
IV.	Conclusiones y trabajos futuros	189
IV.1.	Conclusiones	189
IV.2.	Trabajos futuros.....	194
IV.3.	Publicaciones y conferencias a lo largo de la elaboración de la tesis	195
IV.3.1.	Talleres.....	196
IV.3.2.	Congresos	197
IV.3.3.	Libros y capítulos de libro	199
IV.3.4.	Revistas.....	199
Apéndices	201
	Referencias detalladas de la SLR.....	201
	Localizaciones de las respuestas a preguntas y consecución de objetivos	205
Referencias	209

Listado de figuras

Figura 1 Curva de exageración de Gartner. Fuente: (Gartner Inc, 2016).....	34
Figura 2 Planteamiento metodológico. Fuente: Elaboración propia.	41
Figura 3 Línea de tiempo sobre MOOCs y Open Education. Fuente: (Yuan & Powell, 2013).....	57
Figura 4 Crecimiento de los MOOCs desde el 2012. Fuente: (Shah, 2019).....	59
Figura 5 Dashboard de un estudiante en base al tiempo de dedicación. Fuente: (Amo, Casañ, et al., 2014)	61
Figura 6 Objetivos de uso de los datos en Learning Analytics. Fuente: (Siemens, 2010c)	64
Figura 7 Proceso de Learning Analytics. Fuente: (Chatti et al., 2012).....	65
Figura 8 Diseño de un dispositivo de grabación multimodal para ser utilizado en entornos de lectura: en la sala de clase (izquierda) y desde el punto de vista del estudiante (derecha). Fuente: (Ochoa et al., 2017)	66
Figura 9 Learning Analytics en Google Trends, 01 de junio del 2019.	72
Figura 10 Resumen de términos y artículos en los congresos LAK. Elaboración: propia.	75
Figura 11 Small multiples de los términos más repetidos en los títulos LAK11-19. Fuente propia.	77
Figura 12 Resultado encuesta longitudinal 2018-2019 con relación a el conocimiento leyes de protección de datos en el aula. Elaboración: propia.....	89
Figura 13 Ciclo de exageración de Gartner para negocios que usan tecnologías Blockchain. Fuente: (Gartner, 2019)	93
Figura 14 Organización de DLT, Blockchain y criptomonedas. Fuente: (Dexter, 2018)	96
Figura 15 Pasos y resultados del proceso de revisión y mapeo. Informado como se propone en la declaración PRISMA. Fuente: (Moher et al., 2009)	106
Figura 16 MQ1— Número de artículos publicados por año. Fuente: Elaboración propia	109

Figura 17 MQ3 Canal de publicación de los documentos seleccionados. Fuente: Elaboración propia.....	111
Figura 18 Dominios de aplicación. Fuente: Elaboración propia.....	114
Figura 19 Respuesta a la pregunta PA1 de la encuesta a los estudiantes	153
Figura 20 Respuestas a la pregunta PA2 de la encuesta a los estudiantes	154
Figura 21 Respuestas a la pregunta PA3 de la encuesta a los estudiantes	154
Figura 22 Diagrama de flujo de funcionamiento del primer diseño del plugin "Protected users". Fuente: Elaboración propia.	172
Figura 23 Edición del rol de usuario que permite a un estudiante gestionar el anonimato de sus datos privados. Fuente: Elaboración propia.....	172
Figura 24 Asignación de permisos al rol de usuario que permite a un estudiante gestionar el anonimato de sus datos privados. Fuente: Elaboración propia.....	172
Figura 25 Asignación del rol con permisos de gestión de anonimato al estudiante de un curso en concreto. Fuente: Elaboración propia.	173
Figura 26 El usuario puede gestionar qué datos ocultar desde la cabecera del curso (I). Fuente: Elaboración propia.	173
Figura 27 El usuario puede gestionar qué datos ocultar desde la cabecera del curso (II). Fuente: Elaboración propia.	173
Figura 28 Diagrama de flujo de funcionamiento de la primera evolución del plugin "Protected users". Fuente: Elaboración propia.....	174
Figura 29 Formulario de contacto con el delegado de protección de datos. Fuente: Elaboración propia.....	175
Figura 30 Acceso a protected users como delegado de protección de datos. Fuente: Elaboración propia.....	175
Figura 31 Gestión de los alias de usuarios protegidos versión 1. Fuente: Elaboración propia.....	176
Figura 32 Asignación de alias versión 1. Fuente: Elaboración propia.	176
Figura 33 Acceso a alias desde página del usuario protegido. Fuente: Elaboración propia.....	176

Figura 34 Diagrama de flujo de funcionamiento de la segunda evolución del plugin "Protected users". Fuente: Elaboración propia.....	177
Figura 35 Gestión de peticiones por parte del delegado de protección de datos. Fuente: Elaboración propia.	178
Figura 36 Gestión de alias de usuario versión 2. Fuente: Elaboración propia.	178
Figura 37 Informe administración del alias de un usuario protegido. Fuente: Elaboración propia.....	178
Figura 38 Acceso a alias del usuario protegido desde el perfil de usuario. Fuente: Elaboración propia.....	179
Figura 39 Lista de multialias a los que puede acceder el usuario protegido. Fuente: Elaboración propia.....	179
Figura 40 Diagrama de protección de la tabla de usuarios de Moodle. Elaboración: propia.....	182
Figura 41 Diagrama del plugin Personal Data Broker Log Store. Elaboración: propia.	183
Figura 42 Diagrama de AuthChecker, controlador intermedio entre Moodle y almacenamiento de datos. Fuente: Elaboración propia.	184
Figura 43 Cuadro comparativo de las líneas trabajadas. Elaboración: propia.	185
Figura 44 Resumen de contribuciones y medios. Elaboración: propia.	196

Listado de tablas

Tabla 1 Contextos de estudio de Multimodal Learning Analytics. Fuente: Extraído de (Ochoa et al., 2017)	66
Tabla 2 Diferencias entre Academic Analytics y Learning Analytics. Fuente: (Long & Siemens, 2011)	67
Tabla 3 Utilidades de Social Network Analytics. Fuente: (Amo, García-Peñalvo, & Alier, 2014)	69
Tabla 4 Resultado de los análisis de texto de los artículos LAK11-19	76
Tabla 5 Instrumento de recogida de nivel de conocimiento de leyes educativos por parte de roles educativos de España 2018.....	85
Tabla 6 Instrumento de recogida de nivel de conocimiento de leyes educativas por parte de roles educativos de España 2019.....	86
Tabla 7 Mapeo de preguntas entre cuestionarios del 2018 y 2019.....	87
Tabla 8 Resultados comparativos de las preguntas del cuestionario 2018 y 2019	88
Tabla 9 Cadenas de búsqueda personalizadas para Digital ACM Library.....	103
Tabla 10 Cadenas de búsqueda personalizadas para Web of Science.....	103
Tabla 11 Cadenas de búsqueda personalizada para IEEE Xplore	104
Tabla 12 Cadenas de búsqueda personalizada para Springer Links.....	105
Tabla 13 Lista de control de evaluación de la calidad	108
Tabla 14 Nombres de los autores y número de publicaciones para cada uno	110
Tabla 15 Fuentes de publicación	112
Tabla 16 Resumen de los resultados del informe sobre el mapeo	114
Tabla 17 Taxonomía de trabajos seleccionados en la revisión sistemática de la literatura.....	115
Tabla 18 Resumen de trabajos empleados para responder a las preguntas de investigación.....	115
Tabla 19 Resumen de los resultados de la revisión sistemática	134
Tabla 20 Primera aproximación de encuesta a perfil legal	149
Tabla 21 Preguntas de la encuesta al perfil legal	149

Tabla 22 Respuestas de la encuesta al perfil legal como primera aproximación	150
Tabla 23 Criterios de encuesta a perfiles estudiante	152
Tabla 24 Preguntas de la encuesta a estudiantes	153
Tabla 25 Criterios de entrevista a perfiles profesor	154
Tabla 26 Preguntas de la entrevista a profesores	155
Tabla 27 Respuesta a la entrevista a profesor PP1. ¿El uso de alias complicaría la gestión de aula?.....	155
Tabla 28 Respuesta a la entrevista a profesor PP2. ¿Qué cree que aporta usar un alias como medida de anonimato?.....	157
Tabla 29 Respuesta a la entrevista a profesor PP3. ¿Qué otra solución cree se podría aplicar para resolver el problema del anonimato?	158
Tabla 30 Respuesta a la entrevista a profesor PP4. ¿Ha usado alguna vez un alias para mantener el anonimato?	159
Tabla 31 Resumen de la percepción de cada pregunta de las entrevistas a los profesores.....	161
Tabla 32 Criterios de entrevista a perfiles administradores técnicos de Moodle	162
Tabla 33 Preguntas de la entrevista a los administradores técnicos de Moodle.....	162
Tabla 34 Respuesta a la entrevista a administrador PAT1. ¿Cree que un plugin es una solución técnica compatible con vuestro entorno virtual de aprendizaje?	163
Tabla 35 Respuesta a la entrevista a administrador PAT2. ¿Cree que un plugin de estas características es la solución más adecuada?	163
Tabla 36 Respuesta a la entrevista a administrador PAT3. ¿Qué otra solución complementaria cree que se podría desarrollar?	164
Tabla 37 Respuesta a la entrevista a administrador PAT4. ¿Ha usado alguna vez un alias para mantener el anonimato?.....	164
Tabla 38 Resumen de la percepción de cada pregunta de las entrevistas a administrador técnico de Moodle	165
Tabla 39 Criterios de entrevista a perfiles Delegados de Protección de Datos	166
Tabla 40 Preguntas de la entrevista a delegado de protección de datos	166

Tabla 41 Respuesta a la entrevista a administrador PD1. ¿Crees que eres el perfil adecuado para gestionar las peticiones de anonimato?	167
Tabla 42 Respuesta a la entrevista a administrador PD2. ¿Crees que eres el perfil adecuado para gestionar los alias de usuario?	167
Tabla 43 Respuesta a la entrevista a administrador PD3. ¿Qué otra solución cree que se podría aplicar para resolver el problema del anonimato?	168
Tabla 44 Respuesta a la entrevista a administrador PD4. ¿Ha usado alguna vez un alias para mantener el anonimato?.....	168
Tabla 45 Resumen de la percepción de cada pregunta de las entrevistas a los Delegado de Protección de Datos.....	168
Tabla 46 Resumen de referencias encontradas en la revisión sistemática de la literatura.....	201
Tabla 47 Puntos del manuscrito en los que se responden las preguntas de investigación.....	205
Tabla 48 Puntos del manuscrito en los que se expone la consecución de los objetivos específicos del Objetivos 1	206
Tabla 49 Puntos del manuscrito en los que se expone la consecución de los objetivos específicos del Objetivos 2	206

Glosario

Los siguientes términos son los más utilizados a lo largo de la tesis:

- *Academic Analytics*: Análisis de datos académicos con fines institucionales.
- *Big Data*: Procesamiento de grandes cantidades distinta de datos.
- *Blockchain*: Tecnología de cadena de bloques para evitar el doble gasto.
- *Educational Data Mining (EDM)*: Explotación de datos educativos para extracción de patrones de comportamiento.
- *Learning Analytics*: Analítica del aprendizaje con fines educativos.
- *Learning Management System (LMS)*: Entorno Virtual de Aprendizaje (EVA) como soporte al proceso de aprendizaje en línea.
- *MOOC*: (*Massive Open Online Course* - o Curso Online Masivo y Abierto): Cconsiste en un curso en línea, abierto a cualquiera, sin límite de participantes y de matriculación multitudinaria.
- *RGDP (GDPR)*: El Reglamento General de Protección de Datos (*General Data Protection Regulation*) 2016/679 es un reglamento legislativo de la Unión Europea sobre protección y seguridad de datos personales.
- *Smart Contracts*: Contratos inteligentes para automatizar procesos contractuales.

Siglas y acrónimos

Las siguientes siglas y acrónimos son las más utilizadas a lo largo de la tesis:

- BOE: Boletín Oficial del Estado.
- CEU: Council of European Union.
- COMA: Cursos *Online* Masivos y Abiertos.
- *EDM: Educational Data Mining.*
- *EP: European Parliament.*
- EVA: Entorno Virtual de Aprendizaje.
- DLT: Distributed Ledger Technology
- DPoS: Delegated-proof-of-stake.
- GDPR: General Data Protection Regulation.
- HTTP: HyperText Transfer Protocol.
- HTML: HyperText Markup Language.
- LMS: *Learning Management System.*
- MOOC: *Massive Open Online Course.*
- OA: Objetos de Aprendizaje.
- *OER: Open Educational Resources.*
- PoA: proof-of-accreditation.
- PoB: Proof-of-burn.
- PoD: Proof-of-deposit.
- Pol: Proof-of-importance.
- PoP: Proof-of-personhood.
- PoS: Proof-of-stake.
- PoW: Proof-of-work.
- REA: Recursos de Aprendizaje en Abierto.
- RGDP: Reglamento General de Protección de Datos.

I. Introducción

La contextualización de la presente investigación va íntimamente ligada a la aparición y evolución de las TIC (Tecnologías de la Información y la Comunicación), especialmente en el ámbito de la educación y de forma concreta a los procesos relacionados con los Entornos Virtuales de Aprendizaje (EVA – *Learning Management Systems* (LMS) en inglés) (Adell, Bellver, & Bellver, 2008; Sclater, 2008; Wexler et al., 2007), los Recursos de Aprendizaje en Abierto (REA – *Open Educational Resources* (OER) en inglés; también conocidos de forma genérica como Objetos de Aprendizaje (OA) - *Learning Objects* (LO) en inglés) (D’Antoni, 2012; Glasserman, Mortera, & Montoya, 2013; Ramírez-Montoya, 2015; Ramírez-Montoya & García-Peñalvo, 2015; Salazar Rodríguez, Rodríguez Gómez, & Campos Madrigal, 2012; UNESCO, 2012; Wiley, 2002) y, más recientemente, con los Cursos *Online* Masivos y Abiertos (COMA - *Massive Online Open Courses* (MOOC), en inglés) (Downes, 2012; García-Peñalvo, Fidalgo-Blanco, & Sein-Echaluze, 2018; García Aretio, 2017).

El contexto digitalizado permite a los estudiantes y profesores convertirse en prosumidores (tanto consumidores de contenidos como generadores de ellos). Se crea un ecosistema tecnológico (García-Holgado & García-Peñalvo, 2017; García-Peñalvo, 2018; García-Peñalvo et al., 2017; Llorens, Molina, Compañ, & Satorre, 2014), que a similitud de un ecosistema natural, se define el papel de los organismos como el rol desempeñado por las personas y los componentes de *software*, mientras que el *hardware* permite que el ecosistema funcione. De las interacciones entre todas las partes se desencadena un conjunto de flujos de información. Este ecosistema tecnológico de flujos de información provoca que se genere una cantidad creciente e ingente de todo tipo de datos, generando nuevos ámbitos de investigación como los relacionados con el concepto de *Big Data* (Calvard, 2016; L’Heureux, Grolinger, Elyamany, & Capretz, 2017). Los docentes y administradores de los MOOC se aprovechan de la gran cantidad de datos generados para analizarlos con distintas técnicas estadísticas, matemáticas, basadas en inteligencia artificial, etc. Estas técnicas

se adaptan por los investigadores con conocimientos más pedagógicos creando una aproximación analítica de datos, llamada de forma general, analítica educativa.

El uso de la analítica educativa consiste en la captura y análisis de datos personales, datos y metadatos de las interacciones que los estudiantes realizan en el ecosistema tecnológico, generándose una serie de peligros de privacidad y seguridad. Se distingue en esta analítica educativa dos aproximaciones distintas. El *Academic Analytics* (Campbell, DeBlois, & Oblinger, 2007; Goldstein & Katz, 2005), que se preocupa por mejorar la efectividad organizacional, con alcance hacia arriba institución-gobierno. El *Learning Analytics* (Conde & Hernández-García, 2015; Siemens, 2012, 2013), que se preocupa por mejorar y optimizar el contexto de aprendizaje, con alcance hacia abajo departamento-aula. En la presente investigación se acota el campo de estudio a procesos de *Learning Analytics* debido al bagaje investigador del doctorando, al creciente interés que se deriva del tratamiento de datos educativos, y a la falta de soluciones detectadas a problemas de privacidad en este campo.

Instituciones educativas, investigadores o emprendedores (en adelante, interesados en educación) buscan solucionar los peligros derivados de la aplicación de *Learning Analytics* mediante la implementación de tecnologías emergentes, como *blockchain* (Bartolomé & Lindín, 2019; Bartolomé, Manuel, & Ferrer, 2018; Bartolomé Pina, Bellver Torlà, Castañeda Quintero, & Adell Segura, 2017). *Blockchain* (Swan, 2015) es una tecnología originariamente relacionada con las criptomonedas y orientada a solucionar problemas en el ámbito económico. Por sus posibilidades tecnológicas, e implicaciones sociales, los distintos interesados en educación prometen aplicaciones de la tecnología *blockchain* para solucionar los peligros derivados del uso de *Learning Analytics*. No obstante, el contexto y evolución de las tecnologías en educación es frágil, complejo y no puede solucionarse desde una sola perspectiva tecnológica.

La rápida evolución de las TIC (Berlanga, Peñalvo, & Sloep, 2010; García-Peñalvo, 2015) han transformado y continúan transformado de forma rápida y continua el contexto educativo. La clase magistral ya no es el modelo de aprendizaje único y surgen modelos complementarios mucho más efectivos en los procesos de aprendizaje y de motivación del estudiante (Sein-Echaluze, Fidalgo-Blanco, & García-Peñalvo, 2015). Algunos de

estos modelos, derivados de uso en formación *online*, son incluso antítesis de la clase magistral, como es la Clase Invertida, más comúnmente conocida por su descripción en inglés *Flipped Classroom* (Campanyà, Fonseca, Martí, Amo, & Simón, 2019), y de la que existen múltiples variantes (Fidalgo-Blanco, Martínez-Nuñez, Borrás-Gene, & Sanchez-Medina, 2017; Fidalgo-Blanco, Sein-Echaluce, & García-Peñalvo, 2018; García-Peñalvo, Fidalgo-Blanco, Sein-Echaluce, & Sánchez-Canales, 2019; Sein-Echaluce, Fidalgo-Blanco, & García-Peñalvo, 2019).

Métodos como el comentado, así como modelos de *eLearning* (García-Peñalvo & Seoane Pardo, 2015; Gros & García-Peñalvo, 2016) o *blended learning* (Alier & Casañ, 2008; Graham, 2006) conviven y se entremezclan en formaciones y asignaturas presenciales. La interconectividad e interoperabilidad que ofrecen las TIC (Alier, Casañ, Conde, García-Peñalvo, & Severance, 2010; García-Peñalvo et al., 2015) implican un cambio de paradigma en todos los niveles, ayudando a integrar distintas plataformas, REAs o OAs en un mismo curso académico. Los estándares de intercambio de objetos de contenidos SCORM (Muñoz, García-Peñalvo, Morales, Conde, & Seoane, 2012) facilitan la integración y difusión de contenidos abiertos.

Por otra parte, el uso de las TIC por parte de profesores y estudiantes permite convertirles en creadores de contenidos. Al mismo tiempo, la facilidad con la que plataformas y redes sociales permiten compartir conocimiento en formatos multimedia también facilita el consumo masivo. Este cambio de paradigma convierte, tanto al estudiante como al profesor, en un prosumidor (Islas-Carmona, 2008) de contenidos educativos. En este contexto digital los estudiantes ya no son consumidores de una verdad absoluta, todos pueden crear, comentar y consumir las creaciones de cualquier compañero (Fidalgo-Blanco, Sein-Echaluce, & García Peñalvo, 2017).

A este nuevo paradigma de consumo y producción de contenidos le preceden dos fuerzas de cambio. La primera fuerza es política e induce a la forzada adopción e integración de las TIC en las instituciones educativas (Avanza, 2005; Kraemer, Dedrick, & Sharma, 2009). La segunda tiene relación con la rápida evolución de las TIC en cuanto a dispositivos e infraestructura de conectividad. Las dos fuerzas empujan hacia nuevos contextos como el *Big Data* (Mashey, 1998) en educación (Berendt et al., 2017; Hogan,

2019; Zeide, 2017), en el que se genera una gran cantidad de contenido, de distinto tipo y a gran velocidad.

Las universidades pioneras en ofrecer MOOC (Yuan & Powell, 2013) son las primeras en aprovechar la disponibilidad de grandes cantidades de datos educativos. Los grandes volúmenes de datos educativos se generan a partir de las interacciones de los estudiantes en las plataformas web de cada curso masivo. En este escenario de inscripciones e interacciones masivas, un profesor no es capaz de hacer un seguimiento personalizado y mejorar el aprendizaje de los estudiantes. Estas limitaciones se solucionan aplicando análisis y explotación automatizada de todos los datos educativos recolectados en los MOOC (Brinton & Chiang, 2015; Cruz-Benito, Borrás-Gene, García-Penalvo, Blanco, & Theron, 2017; Romero & Ventura, 2010, 2013), introduciendo un nuevo concepto en la investigación como es la *Educational Data Mining* (EDM), o minería de datos educativos.

La EDM es la ciencia de datos que define modelos, métodos y técnicas de explotación de los datos educativos (Romero & Ventura, 2010). Esta explotación de datos consiste en usar métodos estadísticos, matemáticos y de inteligencia artificial para descubrir y modelizar patrones de comportamiento. El objetivo principal del uso de la EDM es generar información para que los roles de la enseñanza tomen decisiones basadas en datos. Los diseñadores de MOOC terminan integrando todo tipo de técnicas de explotación de datos, algoritmos predictivos y modelado numérico de la realidad educativa para extraer patrones de todo el torrente de datos generado por los estudiantes (Yu, Wu, & Liu, 2019).

En el uso de *Academic Analytics* se analiza y transforma la estrategia de la institución educativa para maximizar los resultados académicos y de rendimiento como organización (Goldstein & Katz, 2005). En el uso de *Learning Analytics* se optimiza y transforma el contexto educativo para que tenga un impacto positivo en el aprendizaje y rendimiento de los estudiantes (Chatti et al., 2012; Long & Siemens, 2011; Next Generation, 2010; Siemens, 2010c).

En los entornos digitales en los que se aplica *Learning Analytics* se recolectan y procesan los datos generados por las interacciones de los estudiantes. La principal técnica para

registrar y extraer datos de las interacciones es el *clickstream* (Amo, Alier, García-Peñalvo, Escudero, & Casañ, 2019; Cody, 1998; Zimmerman, 2001). La aplicación del método *clickstream* en el ámbito educativo consiste en registrar todo el flujo de “clics” que realiza un estudiante en EVA, en páginas web o en *apps* móviles. El flujo de clics permite reconstruir la navegación del estudiante con el objeto de comprender su comportamiento. Para mejorar esta comprensión, los interesados en educación modifican las técnicas y métodos analíticos de *Learning Analytics* para elevarlo a un *Multimodal Learning Analytics* de manera que también se capturen datos de la realidad física del estudiante (Martinez-Maldonado et al., 2016; Ochoa, Lang, & Siemens, 2017; Ochoa, Weibel, Worsley, & Oviatt, 2016; Ochoa & Worsley, 2016). Esta evolución implica un aumento del ecosistema tecnológico, puesto que se añaden nuevos artefactos digitales en el contexto físico de aula, generándose nuevos flujos de información. Mediante la integración del *Modal Learning Analytics* se pueden capturar las emociones faciales o la posición de los estudiantes de forma que se pasa a definir un nuevo objetivo en estos procesos educativos: complementar el *Learning Analytics* en la mejora del contexto de aprendizaje.

Los datos recolectados en procesos definidos por las técnicas y métodos de *Educational Data Mining*, *Academic Analytics* y *Learning Analytics* se almacenan en servidores controlados por las instituciones educativas o por las compañías de los servicios en línea utilizados (Intelliboard, 2019). En cualquier caso, la tendencia es almacenar los datos en centros de datos y de computación, o la llamada nube de computación (en inglés, *cloud computing*). Al-Samarraie y Saeed (2018) apuntan la tendencia en sus palabras el “cloud computing and its applications are vital to the future of distance education worldwide”. En el uso del *cloud computing* los datos trascienden a un nuevo estado donde se convierten en ubicuos, accesibles desde cualquier lugar y fáciles de compartir (Islas-Carmona, 2008). Este nuevo estado adolece de una serie de peligros relacionados con la privacidad, confidencialidad y seguridad de los datos personales de los estudiantes (Robinson, 2017). Cuando estos peligros suceden en entornos educativos en los que también se usa *Learning Analytics*, entonces se generan situaciones de desconfianza y

aparecen recelos y dudas (Drachsler, 2016), ya que los datos pueden ser usados, compartidos y explotados por y entre agentes (Herold, 2014).

La confianza en el uso de datos requiere de un adecuado nivel de seguridad y privacidad en las comunicaciones digitales. Si las transmisiones no son seguras tampoco va a ser seguro enviar información sensible por Internet, por ejemplo, los datos personales, académicos, demográficos o económicos. La evolución de los estándares de seguridad añade capas de confianza a las transacciones digitales, pero no obstante, aún quedan inconsistencias tecnológicas sobre privacidad que son cuestiones que se ponen de manifiesto en los inicios de Internet y que de alguna manera aún perduran (Farrell & Tschofenig, 2014; Schneier & Hardie, 2014; Tschofenig & Baccelli, 2019).

En una fecha tan temprana como 1981 ya había expertos tratando de resolver con criptografía los problemas de privacidad, seguridad e inclusión que internet planteaba (Tapscott & Tapscott, 2016).

Satoshi Nakamoto (2008) crea la criptomoneda *Bitcoin* y difunde su propuesta tecnológica subyacente. Satoshi Nakamoto es en realidad un seudónimo del creador, cuya identidad aún permanece en el anonimato. El gran interés de esta nueva criptomoneda subyace en los posibles usos de la tecnología propuesta, de hecho, se ha generado mucha expectativa y grandes promesas para solucionar problemas que requieran confianza digital en entornos de poca confianza.

Nakamoto propone un protocolo de encadenamiento de bloques de transacciones económicas (del inglés *blockchain*), cuyo cometido principal es solucionar el problema del doble gasto: evitar que se pueda gastar una moneda virtual dos o más veces. Una moneda física solo puede usarse una sola vez, puesto que el emisor se desprende de la moneda físicamente al entregarla al emisor. En cambio, una moneda virtual puede usarse infinitas veces, puesto que solo existe en lo digital, entorno en el que cualquier representación puede copiarse y reproducirse infinitamente. Para que no se pueda realizar un doble o más gasto de una moneda virtual, debe haber una entidad central

que controle la emisión y recepción de monedas. Con la propuesta tecnológica de Nakamoto se soluciona el problema del doble gasto sin la necesidad de una entidad central, por tanto, pueden evitarse intermediarios en transacciones económicas y generarse un contexto de confianza en entornos de baja confianza.

Nakamoto propone un protocolo de almacenamiento descentralizado de bloques de transacciones encadenados y validados por consenso mediante pruebas (en adelante tecnología *blockchain*). En el protocolo original las pruebas de trabajo son computacionales y criptográficas, más adelante se proponen, por otros autores, otros tipos de pruebas más adecuadas al contexto de aplicación, por ejemplo, *proof-of-stake* (PoS), *delegated-proof-of-stake* (DPoS), *proof-of-importance* (PoI), *proof-of-personhood* (PoP), *proof-of-burn* (PoB), *proof-of-deposit* (PoD), *proof-of-accreditation* (PoA) (Aste, Tasca, & Di Matteo, 2017; Borge et al., 2017; Duan, Zhong, & Liu, 2018; Zheng, Xie, Dai, Chen, & Wang, 2017). Para conseguir el consenso se propone construir una red descentralizada de nodos computacionales que puedan realizar las pruebas de trabajo. Los usuarios que gestionan los nodos que consiguen terminar las pruebas de trabajo reciben una cantidad de la criptomoneda a modo de compensación por el esfuerzo (Nakamoto, 2008). El uso de pruebas de trabajo computacional implica que al quererse realizar un cambio en algún bloque de transacciones de la tecnología *blockchain*, se deban volver a realizar todas las pruebas de trabajo para re-encadenar los bloques posteriores al bloque modificado, hecho que supone un trabajo computacional muy elevado. Este modo de operar confiere inmutabilidad a los datos almacenados en los bloques de transacciones, característica fundamental para conferir seguridad transaccional. Todos los nodos/actores tienen una copia de los registros de transacciones, por este motivo un sistema que utilice tecnología *blockchain* se considera como un sistema descentralizado y transparente.

Las características de inmutabilidad, transparencia, descentralización y pruebas de consenso solucionan en conjunto el problema del doble gasto y genera un contexto de confianza entre los nodos de la red. Se crea así una economía de la confianza (Botsman, 2012) capaz de cambiar la manera de hacer negocios e incluso de trasladar al ciudadano digital el control de sus datos (Molins, 2019).

La nueva economía de la confianza supone una fuerte revolución en todos los ámbitos de la sociedad, sobretodo cuando la confianza se automatiza, pasando de una confianza basada en intermediarios a una de automatizada por código bautizada como *Smart Contracts* (Szabo, 1997). Un *Smart Contract* es un pequeño programa que funciona a modo de contrato inteligente que ejecuta automáticamente acuerdos contractuales al cumplirse ciertas condiciones y permite eliminar intermediarios en cualquier transacción entre pares. Su desarrollo en *Bitcoin* ofrece capacidades limitadas, en cambio, en otras plataformas como *Ethereum*, se pueden desarrollar con un lenguaje de programación de Turing completo para crear programas capaces de realizar cualquier cálculo. Por consiguiente, es posible automatizar acuerdos contractuales dentro de las soluciones desarrolladas con tecnología *blockchain* para que sucedan eventos transaccionales, por ejemplo, la entrega de un título académico cuando se ha conseguido superar un curso.

La evolución de la adopción de la tecnología *blockchain* viene explicada por la curva del *hype cycle* de Gartner (Gartner Inc, 2016), un modelo de referencia que define las fases de adopción de una tecnología (ver Figura 1). Su objeto es descubrir su potencial de solvencia ante oportunidades y problemas, inclusive en educación (Prinsloo & Van Deventer, 2017).

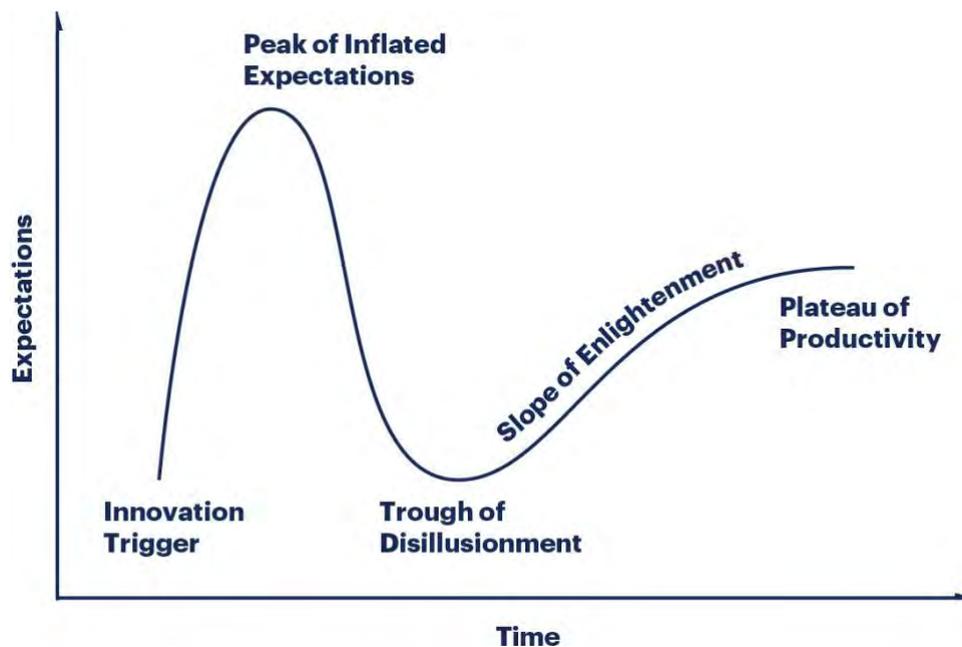


Figura 1 Curva de exageración de Gartner. Fuente: (Gartner Inc, 2016)

Gartner analiza la adopción de la tecnología *blockchain* desde su creación. En el 2019 lo sitúa en el abismo de la desilusión (Litan & Leow, 2019). En educación esto significa que se empiezan a entender las posibilidades reales de los usos de la tecnología *blockchain*, qué carencias presentan estos posibles usos y qué problemas reales podrían solucionar (Poveda, 2018).

Como se comprueba de la contextualización realizada, se han ido generando en las últimas décadas conexiones entre métodos, tecnologías y sistemas educativos con el tratamiento, gestión y seguridad de los datos educativos, que merecen de un estudio detallado sobre el alcance resolutivo de la tecnología *blockchain* en cuestiones de privacidad y seguridad dentro de procesos de *Learning Analytics*.

I.1. Acotación del objeto de estudio

La recolección y tratamiento de los datos educativos en procesos de *Learning Analytics* enmarcan un contexto complejo. Cada una de las herramientas usadas en el aula registra las interacciones de los estudiantes en bases de datos de servidores remotos (Miguel Á. Conde et al., 2014; Gómez-Aguilar, García-Peñalvo, & Therón, 2014; Gómez-Aguilar, Hernández-García, García-Peñalvo, & Therón, 2015). Datos personales, datos académicos y metadatos de estudiantes se almacenan dispersos y fuera del control de los propios estudiantes. Los estudiantes quedan expuestos ante cualquier entidad tercera que capture sus interacciones educativas (Bartolomé Pina et al., 2017; Singer, 2014; Williamson, 2017b). Por añadidura, en la mayoría de las herramientas ni las mismas instituciones educativas tiene voz alguna en la gestión y control de los datos recolectados de sus estudiantes.

Asimismo, puede que la información recolectada sea modificada, transformada y compartida con terceros sin el consentimiento oportuno (Lupton & Williamson, 2017; Robinson, 2017). Por consiguiente, los estudiantes interactúan con herramientas de aprendizaje en un entorno altamente frágil en cuestiones de confidencialidad de datos e identidad de los estudiantes (Pardo & Siemens, 2014), problemas que de forma resumida podríamos identificar como que:

- **Existe una marcada dispersión de datos.** Cuantas más herramientas se utilizan en el aula, los datos de los estudiantes están más dispersos y corren más riesgo de ser mal utilizados.
- **La protección de los datos es baja.** Los datos son almacenados en abierto, los alumnos se identifican claramente (Singer, 2014), pueden ser modificados a conveniencia (Patents, 2018) y accedidos libremente por terceros e incluso compartidos sin consentimiento fuera del amparo de la ley.
- **No hay ningún control sobre los datos.** Los estudiantes no pueden determinar quién, cuándo y durante cuánto tiempo se puede acceder a los datos que generan en las herramientas educativas.

El tema objeto de investigación se centra en una problemática común en cualquier contexto educativo: la fragilidad en cuanto a la protección de la confidencialidad de datos e identidad de los estudiantes en el uso de *Learning Analytics* (Drachsler & Greller, 2016; Manel & Sanz, 2016; Pardo & Siemens, 2014; Siemens, 2012), especialmente en sistemas de aprendizaje mediados por EVA (García-Peñalvo, 2005, 2008; García-Peñalvo & Seoane-Pardo, 2015).

A esta fragilidad se pone de manifiesto la necesidad de una regulación legal en la recolección y tratamiento de datos personales (EP and the CEU, 2016). Las leyes definen un marco legal que permite castigar a los infractores, aunque se encuentran estudios donde se comprueba que no es suficiente para terminar con usos indebidos (Herold, 2014).

Ante el problema de la fragilidad se requiere ceder al estudiante el poder de decisión de quién, cuándo, cuánto tiempo y a qué datos tendrá acceso. Esta idea se resume en otorgar el control de los datos generados por el estudiante al propio estudiante. Este contexto de control se consigue haciendo cumplir la legalidad de manera que los estudiantes puedan cumplir con sus derechos de protección de datos.

En esta investigación se apunta y pone en estudio a la tecnología *blockchain* (Nakamoto, 2008) como complemento a proceso de *Learning Analytics* para dar solución al problema. La tecnología *blockchain* se acogió entre grandes expectativas, muchas promesas y fuertes críticas (Bartolomé Pina et al., 2017), de forma que se puede

considerar como una candidata para solucionar el problema a la situación de fragilidad en cuanto a sus cualidades de descentralización, distribución, inmutabilidad, anonimidad y confianza por consenso. No obstante, y llegados a este punto, cabe preguntarse, desde la perspectiva tecnológica, ¿cuáles son las verdaderas oportunidades de la tecnología *blockchain* para educación?, por un lado, y ¿pueden implementarse soluciones con la tecnología *blockchain* que resuelvan el problema?, por otro.

Además, en una red de redes construida encima de unas tecnologías originariamente abiertas y no orientadas a la seguridad y privacidad de los cibernautas, incluidos los estudiantes ¿es la tecnología *blockchain*, en su concepción original, capaz de asegurar un nivel adecuado de confidencialidad de datos en procesos de *Learning Analytics*?

La investigación destaca la existencia de unas leyes marco definitorias de los derechos de los usuarios en el tratamiento de sus datos personales. En este sentido, cabe preguntarse si las soluciones en educación que solo usan la tecnología *blockchain* como almacén de datos cumplen con la legalidad o tales aplicaciones deberán complementarse con tecnologías híbridas de base tradicional y centralizadas.

La presente tesis nace con la motivación de dar respuesta a las preguntas anteriores, más las surgidas a medida que la propia investigación avanza. Como se verá en el siguiente apartado, alguna de estas nuevas preguntas ha desembocado a su vez en la necesidad de diseñar, implementar y validar nuevos prototipos funcionales a carencias de las que adolece el ámbito educativo en el plano de la confidencialidad de datos y/o seguridad de estos.

I.2. Objetivos y preguntas de investigación

La unión de la capa tecnológica *blockchain* al *Learning Analytics* como solución al problema es un área de estudio en progreso y a la vez compleja, tal y como se justifica en la revisión sistemática realizada. Proteger los datos personales de los estudiantes presenta una serie de desafíos y retos que requieren de nuevas perspectivas tecnológicas, capas de políticas de seguridad y nuevos procesos de adecuación legal. Superar la problemática mediante nuevas capas tecnológicas de seguridad parece ser el

procedimiento lógico, pero ¿puede ser *blockchain* una tecnología que en su aplicación permita resolver el problema detectado? Esta investigación pretende estudiar las características y posibilidades de la tecnología *blockchain* para ser implementada en una posible solución a este problema, así como estudiar en profundidad las diversas soluciones actuales que ya la implementan con el fin de proteger la privacidad e identidad de los estudiantes, de forma que los resultados puedan transferirse en la recolección y tratamiento de datos educativos en procesos de *Learning Analytics*.

I.2.1. Objetivos

La presente tesis tiene como objeto de investigación la privacidad, la confiabilidad y la seguridad de los datos relacionados con la actividad de los actores involucrados en un proceso de enseñanza/aprendizaje mediado un ecosistema tecnológico de aprendizaje en el que se incluyen servicios de *Learning Analytics*. Para ello se persiguen dos objetivos generales:

- O.1. Demostrar que el *Blockchain* es una tecnología que puede aportar una posible solución viable al problema de la falta de privacidad, confidencialidad y seguridad de los datos recolectados y usados en procesos de *Learning Analytics* en ecosistemas tecnológicos de aprendizaje.
- O.2. Diseñar e implementar una solución tecnológica para adecuar el nivel de confidencialidad de datos personales educativos a lo impuesto por el Reglamento General de Protección de Datos (RGPD) cuando se tienen procesos de *Learning Analytics* en ecosistemas tecnológicos de aprendizaje.

La consecución de los objetivos de la investigación parte de un esfuerzo por conocer las leyes que otorgan derechos a la sociedad. Tales derechos deben estar reflejados tanto en el desarrollo como disponible su ejercicio en el uso de las herramientas educativas. Relacionar las leyes con la tecnología como objeto de validación es importante para poder establecer un marco estándar sobre privacidad y seguridad. De esta comprensión

y delimitación de la realidad tecnológico-legal, parten los siguientes objetivos específicos.

Los objetivos específicos asociados al objetivo O.1 son:

- OE.1.1. Comprender la necesidad de utilizar una tecnología protectora de la privacidad e identidad de los estudiantes.
- OE.1.2. Validar el uso de la tecnología *blockchain* como protectora de la privacidad e identidad de los estudiantes en procesos de *Learning Analytics*.
- OE.1.3. Desarrollo de un prototipo que implemente la tecnología *blockchain* para proteger la privacidad e identidad de los estudiantes en el uso de *Learning Analytics*.
- OE.1.4. Integrar el uso de *Smart Contracts* en la tecnología *blockchain* como marco de políticas de gestión de datos con terceros.

Los objetivos específicos asociados al objetivo O.2 son:

- OE.2.1. Estudiar el impacto del Reglamento General de Protección de Datos en la aplicación de *Learning Analytics*.
- OE.2.2. Estudiar el impacto del Reglamento General de Protección de Datos en los ecosistemas tecnológicos de aprendizaje.
- OE.2.3. Detectar carencias en los ecosistemas tecnológicos de aprendizaje para la correcta implantación del Reglamento General de Protección de Datos.
- OE.2.4. Desarrollo de un prototipo en un Entorno Virtual de Aprendizaje, como un de los componentes principales de un ecosistema tecnológico de aprendizaje, que asegure un nivel adecuado de confidencialidad y seguridad de datos personales del alumno impuesto por el Reglamento General de Protección de Datos.

En consecuencia, esta investigación tiene una clara doble índole. Por una parte, se centra en validar la tecnología emergente *blockchain* como tecnología útil para preservar la privacidad y la seguridad de los datos personales educativos, para finalmente crear un prototipo funcional. Por otra parte, la investigación se desvía a otro

ámbito, aún mucho más legal, que permite comprender, desde otro ángulo distinto al tecnológico, las connotaciones y el alcance de las palabras privacidad y seguridad.

I.2.2. Preguntas de investigación

Para conseguir el grado de compleción de los objetivos generales propuestos, se define una serie de preguntas de investigación:

- P.1. ¿Cuáles son los orígenes y evolución del análisis de datos educativos?
- P.2. ¿Cuáles son los problemas relacionados con el análisis de datos educativos?
- P.3. ¿Qué implicaciones tiene asegurar la privacidad y seguridad de los datos educativos en procesos de *Learning Analytics*?
- P.4. ¿Cuáles son las soluciones que implementan la tecnología *blockchain* para resolver problemas en el contexto educativo?
- P.5. ¿Existe alguna aplicación concreta de la tecnología *blockchain* que pueda considerarse como una solución al problema planteado y cumplir con las leyes de protección de datos?
- P.6. ¿Puede solucionarse el problema con una implementación de la tecnología *blockchain* sin la ayuda de otras tecnologías de almacenamiento de datos?
- P.7. ¿Puede asegurarse un adecuado nivel de protección de datos en entornos de aprendizaje sin necesidad de usar tecnología *blockchain*?

I.2.3. Metodología

Para lograr los objetivos propuestos se plantea un marco metodológico multi-método con la voluntad final de desarrollar, validar y evaluar un prototipo en base a un proceso iterativo e incremental, con las fases de análisis, diseño, implementación y evaluación (Pressman & Maxim, 2015). La Figura 2 muestra el esquema del planteamiento metodológico utilizado.

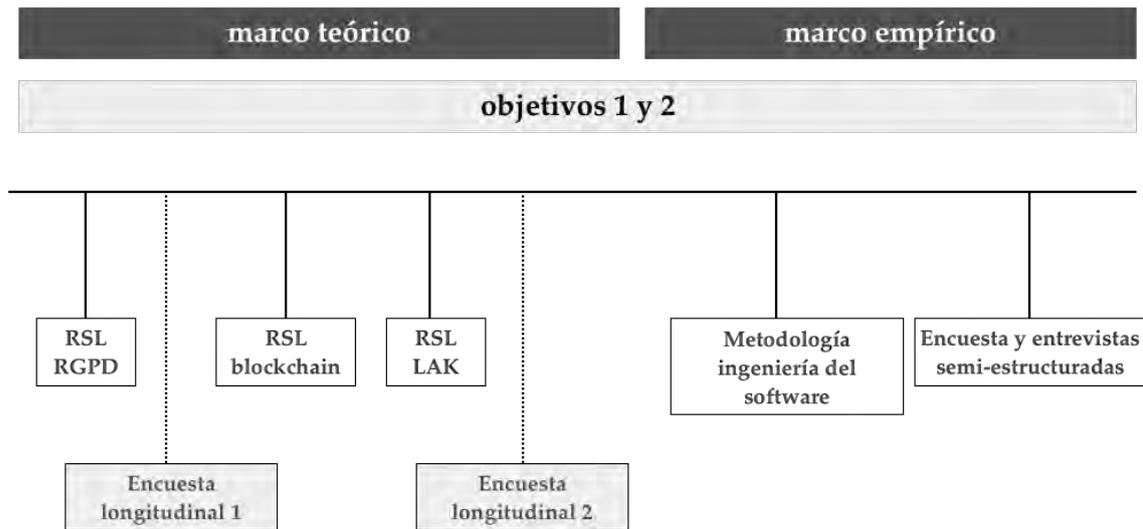


Figura 2 Planteamiento metodológico. Fuente: Elaboración propia.

El planteamiento metodológico se compone de los siguientes métodos, relacionados con los objetivos generales:

- O.1. Tiene una estrecha relación con el marco teórico que sustenta la investigación:
 - Revisión de la literatura. Las revisiones de la literatura de los distintos contextos (literatura, debate, etc.) son la manera que los investigadores disponen para construir un estado de la cuestión, conocer cuáles son las investigaciones ya realizadas, comprender qué debate discurre en el ámbito de estudio, unir puntos de información, encontrar paralelismos entre investigaciones o conocer nuevos enfoques metodológicos. En definitiva, cualquier revisión relevante al problema permite al investigador adquirir conocimiento profundo del área a investigar para progresar con paso firme en las siguientes fases. Concretamente, se realizan tres revisiones desde perspectivas distintas:
 - Revisión sistemática de literatura en el ámbito de la tecnología *blockchain* y educación para analizar y sintetizar el estado de la cuestión (García-Peñalvo, 2019; Kitchenham et al., 2009; Kitchenham & Charters, 2007). La revisión sistemática de la literatura llevada a cabo tiene el objetivo de conocer el estado

de la cuestión de la tecnología *blockchain* en su uso en educación, las soluciones existentes ligadas a la problemática y analizar procedimientos ejecutados por otros investigadores. La revisión sistemática para esta investigación comprende distintas áreas:

- Las experiencias, usos y teorías sobre la tecnología *blockchain* que aborden soluciones a problemas en educación.
 - Las aplicaciones, metodologías, modelos, estrategias, aproximaciones y adopciones de *Learning Analytics* con la tecnología *blockchain*.
 - Automatizaciones dentro de la tecnología *blockchain* realizadas con *Smart Contracts*.
 - Problemas o incompatibilidades de privacidad en la implementación de la tecnología *blockchain* tanto en el campo de la ciberseguridad como de las leyes de protección de datos personales.
- Revisión del contexto legal. A modo de conocimiento más profundo del estado de la cuestión, se realiza una revisión de las distintas leyes vigentes a nivel estatal, europeo e internacional. El conocimiento extraído da pie a una encuesta longitudinal, de inicio 2018 y término 2019, para conocer el nivel de adopción de las leyes de protección de datos en roles educativos de España.
 - Revisión de congresos *Learning Analytics and Knowledge (LAK)*. Como tercera aproximación y ampliación del conocimiento sobre el estado del arte se realiza una revisión de la literatura en el ámbito de *Learning Analytics* y aspectos éticos, de privacidad y de seguridad. Esta revisión se puntualiza en el contexto de los congresos LAK porque se les considera de

referencia obligada en el ámbito de análisis de datos educativos, tal y como se verá en la sección II.5.1 Congresos Learning Analytics and Knowledge.

- Entrevistas semiestructuradas. Con un enfoque cualitativo y con el objetivo de completar el conocimiento del contexto legal, se quiere conocer su uso entre educadores de España. A modo de conocimiento más profundo del estado de la cuestión se realizan dos series de entrevistas semiestructuradas para conocer el nivel de adopción de las leyes de protección de datos en roles educativos de España. Una serie de entrevistas se realiza en septiembre del 2018, etapa en la que se hace efectivo el RGPD, y otra serie en septiembre del 2019. Conocer cómo los roles educativos hacen uso de estas leyes, sus conocimientos al respecto, cómo actúan los centros para transferir conocimientos prácticos que impacten en el aula y la preocupación en general por la privacidad y seguridad de los involucrados en el proceso de enseñanza/aprendizaje en el tratamiento de sus datos, se entiende como esencial para comprender la magnitud del problema.
- O.2. Se utiliza el resultado de la revisión de la literatura para definir los requisitos relacionados con el desarrollo del prototipo funcional, propio del O.2, con tecnología *blockchain* que supere las posibles deficiencias encontradas a lo largo de la triple revisión de los artículos, que adopte metodologías de otros investigadores ya desarrolladas y útiles para la consecución del objetivo. Se aborda este segundo objetivo con la motivación de procurar un nivel adecuado de confidencialidad de datos y seguridad establecido por las leyes de protección de datos en ecosistemas tecnológicos de aprendizaje. Para conseguir este segundo objetivo se sigue investigando en entornos educativos que usan análisis de datos de los actores involucrados, pero apartando del camino cualquier aspecto relacionado con la tecnología *blockchain*.

- Entrevistas semiestructuradas. De nuevo con un enfoque cualitativo, se realizan una serie de entrevistas con actores del contexto educativo (profesores, directores, estudiantes de TFM), negocios (desarrolladores, *partners* de Moodle) o investigación (catedráticos, doctorados, doctorandos) relacionados con conceptos como *blockchain*, *Learning Analytics*, *Big Data*, computación, seguridad, privacidad, leyes o automatización. El conocimiento obtenido de tales personas ha aportado nuevos enfoques e ideas que han reforzado la disertación y aportación científica de la presente tesis. Las entrevistas se formulan en base a una metodología de experiencia de usuario cuyos resultados fundamentan los requisitos para el desarrollo del prototipo.
- Método de ingeniería de software. Se desarrolla, mediante un proceso iterativo e incremental, un prototipo funcional que permita ejercer ciertos derechos legales específicos no disponibles actualmente en un entorno virtual de aprendizaje concreto. Por cuestiones éticas, el prototipo utiliza datos de registros despersonalizados para evitar posibles conflictos éticos y legales relacionados con las leyes de protección de datos vigentes.

I.3. Marco de trabajo

Este trabajo se desarrolla en el Programa de Doctorado Formación en la Sociedad del Conocimiento de la Universidad de Salamanca (García-Peñalvo, 2013, 2014, 2019b). El programa está regulado por el Real Decreto 99/2011 (Gobierno de España, 2011). El trabajo es a la vez codirigido por dos directores adicionales, uno adscrito a la Universitat Politècnica de Catalunya y el otro adscrito a La Salle Barcelona – Universitat Ramon Llull.

Para la compleción de los objetivos marcados en esta tesis doctoral se dispone del soporte del grupo de investigación GRIAL (García-Peñalvo, 2016b; Grupo GRIAL, 2019; Peñalvo et al., 2012) dirigido por Francisco José García Peñalvo, uno de los directores de esta tesis, del grupo de investigación GRETEL al que pertenezco y dirigido por David

Fonseca, director de esta tesis, y del grupo de investigación BCN SEER, al que pertenece Marc Alier, director también de esta tesis.

Los tres grupos tienen líneas de investigación relacionadas con *Learning Analytics*, educación, sociedad del conocimiento y Big Data, conceptos directamente relacionados con la propuesta de tesis. Los recursos humanos disponibles para debatir estas distintas cuestiones son de gran ayuda para dilucidar y establecer caminos dentro de la propia investigación.

I.4. Organización del documento

El documento se organiza en cuatro capítulos a lo largo de los cuales se discurre y reflexiona sobre la investigación desarrollada. En el primero, Introducción, se muestra el marco definitorio de la investigación. Este capítulo está compuesto por cuatro partes, en las cuales se realiza una primera aproximación al estado del arte y cómo se abordará científicamente.

El segundo capítulo Marco teórico, profundiza en el estado del arte y se divide en ocho partes. Se clarifican conceptos y términos fundamentales para la comprensión del trabajo. A la vez se presentan acciones llevadas a cabo, como una revisión sistemática de la literatura o una encuesta a educadores, para terminar de profundizar en el estado del arte y aportar información actualizada a la literatura consultada.

En el tercer capítulo Marco empírico se muestra la aproximación más científico-práctica de la investigación y se divide en cinco partes. Se presentan las fases relacionadas con el marco empírico y sus distintas fases, como los desarrollos y prototipos funcionales fundamentados en el marco teórico.

En el capítulo cuatro se concluyen los resultados de esta investigación y se muestran futuras líneas de investigación, así como las publicaciones realizadas durante la misma. Como cierre del documento se encuentran los apéndices y el compendio de todas las referencias usadas a lo largo de la investigación.

II. Marco teórico

En este capítulo se expone el contexto teórico del problema abordado en esta tesis en toda su magnitud. La exposición transcurre en los distintos estadios evolutivos de la tecnología, su impacto en educación y consideraciones tanto de privacidad como de seguridad. Por consiguiente, para situar el contexto en el que desemboca el problema se considera necesario introducir:

- Una aclaración de los conceptos privacidad, seguridad y confidencialidad de datos personales.
- El origen la tecnología de red que interconecta a los involucrados en el proceso de enseñanza/aprendizaje y su relación con la seguridad y privacidad de sus datos.
- Los MOOC como puerta de entrada del *Big Data* en educación.
- La coyuntura *Big Data* en educación.
- Cuestiones sobre privacidad en el tratamiento de datos educativos en procesos de *Learning Analytics*.
- Las leyes y marcos legales que reglamentan y regulan el tratamiento de datos personales en el uso de las TIC.
- *Blockchain* como tecnología emergente, su relación con educación y promesas para resolver los problemas.

En la exposición se ahonda en las definiciones, los conceptos, los términos, los marcos y los principios para facilitar la comprensión de la problemática a la que hace referencia la presente investigación. Una vez presentado el marco teórico se entra a mayor profundidad en el estado de la cuestión mediante un proceso de revisión sistemática de la literatura.

II.1. Concepto de privacidad

Los conceptos de privacidad y confidencialidad se han confundido y se están utilizando históricamente de forma incorrecta (Francis, 2008). Por dichos motivos, es necesaria una apreciación en los conceptos de privacidad y confidencialidad de datos personales,

exponer el contexto en el que se debe usar una palabra u otra y clarificar qué usos se le da a lo largo del presente trabajo.

La palabra privacidad no se utiliza de forma correcta (Francis, 2008). En contextos en los que la confidencialidad de datos es el concepto adecuado a aplicar, se usa privacidad de forma errónea. Esto es debido a una adopción anglosajona del término *privacy*. Es necesario aclarar conceptos y exponer la manera en que se utilizan en esta investigación, para evitar los usos incorrectos generalizados. Por consiguiente, y a continuación, se expone la diferencia entre privacidad y confidencialidad, qué concepto se usará a lo largo de la investigación y cómo afecta a cualquier búsqueda de trabajos y referencias relativas al problema.

La palabra seguridad no lleva a confusión. Se deja como último punto para exponer su uso en la presente investigación.

II.1.1. Privacidad o confidencialidad de datos personales

La palabra privacidad hace referencia a proteger cuestiones física e íntimas de una persona, como por ejemplo protegerla de tocamientos corporales. Por el contrario, la palabra confidencialidad hace referencia a la salvaguarda en la transmisión de información, referidas a acciones tales como compartir datos académicos o registros sanitarios. En resumen, la privacidad es invadida y la confidencialidad es incumplida (Francis, 2008; Santiago, 2016).

La palabra privacidad proviene de la anglosajona *privacy* y su traducción a distintos idiomas de la unión europea hace referencia a la intimidad física (Francis, 2008). En Estados Unidos, la palabra privacidad va muy ligada a la protección de registros médicos y, por tanto, a cuestiones íntimas. Es por este motivo que se le atribuye un significado de protección de la información personal, motivo por el cual se ve extendido su uso en contextos en el que se hace referencia a protección de datos (Francis, 2008).

En cambio, las leyes de protección de datos europeas hacen referencia a la salvaguarda de la información personal recolectada, y, por ello, usan confidencialidad de datos. Ejemplo de ello es el texto legal del RGPD, en el que se utiliza confidencialidad en el ejercicio de la protección de datos en lugar de privacidad (EP and the CEU, 2016).

Por cuestiones de intimidad se establece un orden de importancia, donde la privacidad se sitúa por encima de la confidencialidad (Francis, 2008). Sin embargo, en las leyes de protección de datos la palabra adecuada es confidencialidad, puesto que el contexto legal hace referencia a la gestión de información de los datos de una persona. En otros contextos, sobretodo en línea, la palabra privacidad también se refiere a la protección de la identidad de las personas. En consecuencia, depende del contexto el utilizar palabras como privacidad, protección de la identidad o confidencialidad de datos personales.

En la presente investigación, en el contexto de tratamiento de datos personales, se debería usar la palabra confidencialidad en lugar de privacidad. El porqué de esta decisión se debe en distintos motivos:

- El tratamiento de datos personales en procesos de *Learning Analytics*.
- El investigador forma parte de un país del territorio europeo donde privacidad hace referencia a la intimidad física.
- La palabra confidencialidad hacer referencia a un tratamiento de datos personales.
- Solo aparece confidencialidad en los textos legales del RGPD, inclusive en los textos en inglés.

No obstante, a pesar de los motivos expresados, se decide usar la palabra privacidad para seguir una coherencia con todos los trabajos consultados en la revisión sistemática, puesto que el idioma de todos estos es el inglés, y evitar confusiones o apreciaciones constantes en el texto. Al mismo tiempo, se procura contextualizar el uso de la palabra privacidad en cuanto se refiera a confidencialidad de datos, permisos y roles de acceso a datos o preservación del anonimato de la identidad *online*.

II.1.2. Seguridad

La seguridad, en entornos digitales, implica proteger de ataques y en caso de acceso indebido, evitar la comprensión de los datos obtenidos. Los mecanismos de encriptación son esenciales para asegurar la seguridad dictaminada en la legislación sobre protección de datos personales (EP and the CEU, 2016). Estas leyes establecen unos niveles

adecuados de seguridad de datos ante el tratamiento de datos personales. Por desgracia, las leyes de protección de datos surgen para proteger a los ciudadanos de herramientas digitales que hacen un mal uso de las trazas recolectadas, incluso en contextos educativos (Isaak & Hanna, 2018; Robinson, 2017).

Para comprender la sensibilidad del contexto educativo en la que se enmarca la investigación, se requiere exponer la infraestructura de privacidad abierta que da soporte a los servicios disponibles en Internet:

- Primero, para que el lector comprenda los orígenes y evolución del *Learning Analytics*.
- Segundo, para asimilar la importancia y valor de los datos en la coyuntura actual de vigilancia y capitalismo de datos (West, 2019; Zuboff, 2015).
- Tercero, para asentar los abusos que se están llevando a cabo en cuestiones de privacidad de datos en el contexto educativo y el interés creciente de las inversiones empresariales en educación.
- Y cuarto, para exponer posibles soluciones al problema.

II.2. Internet insegura y *clickstream*

La red de redes socava la privacidad de los cibernautas del siglo XXI y, por ende, la de los estudiantes, cuyo contexto educativo puede ser gobernado por empresas tecnológicas mediante el análisis de sus datos recolectados (Williamson, 2016a, 2016b).

Los cimientos protocolarios de Internet hasta los comportamientos de uso de los cibernautas actuales, inclusive estudiantes, dibujan una realidad digital de vigilancia activa (Zuboff, 2015). Es necesario preservar la privacidad *online* de los participantes en un proceso de enseñanza/aprendizaje a través de Internet en todos los sentidos para evitar que se conviertan en materias primas del capitalismo de datos (West, 2019).

En esta situación (distópica) de vigilancia activa se hace muy fácil hacer un seguimiento constante de las personas que transitan en el mundo digital, incluido el educativo. Dicho seguimiento desemboca en una educación gobernada por datos, con sus peligros y consecuencias (B. K. Daniel, 2019), en el que se otorga al *Big Data* la calidad de verdad

absoluta (Hogan, 2019). Esta situación, implica que en esta práctica de gobernanza digital subyagan prácticas de ciencia de datos, métodos y maneras de pensar que apelan a una objetividad técnica y neutral de observaciones empíricas y análisis. Por consiguiente, se afirma que con el *Big Data* se podría llegar a comprender cómo se enseña y cómo leen los estudiantes (Williamson, 2017a). No obstante, el uso de datos educativos recolectados en forma masiva también se usa como arma de gobernanza educativa (Williamson, 2017a). Se requiere traspasar la gobernanza al estudiante para preservar los niveles de confidencialidad y seguridad de datos establecidos en los marcos legales vigentes (Boletín Oficial del Estado, 2018; EP and the CEU, 2016).

II.2.1. Estadios en la evolución de internet

Las TIC se sustentan en una evolución de cuatro estadios, cada uno construido sobre el anterior (Baricco, 2019). Estos estadios hacen posible una captura de datos constante y usos indebidos que explican el socavamiento de la privacidad cibernauta, y son: Internet, Web, plataformas y *apps*. El origen de esta distopía está en el capitalismo fragmentado, que en palabras del filósofo Francesc Llorens:

La crisis de la privacidad es un efecto directo de la fragmentación del capitalismo, originalmente un fenómeno económico, en "formas" de mayor concreción: capitalismo emocional, cognitivo, de plataformas, de vigilancia...

(Llorens, comunicación privada por Twitter, 3 de septiembre del 2019)

El capitalismo de plataformas y vigilancia crea el problema de privacidad que se remarca en la investigación y del que adolece el contexto educativo. Los datos educativos se explotan y se entienden como materia prima que refinar para conseguir rédito. Por consiguiente, el contexto educativo es sensible y requiere de una seguridad y privacidad de datos adecuada a las leyes de protección de datos, en concreto, al RGPD (EP and the CEU, 2016).

II.2.1.1. Internet

El protocolo TCP/IP (Stevens, 1994), que activa la circulación por las autopistas digitales de la información, término acuñado por Al Gore (Broad, 1992), no es privado. Los protocolos de transmisión permiten enviar paquetes de datos a las redes informáticas. Estos paquetes están diseñados por defecto para contener dos partes; una con la información de la transmisión y la otra con el contenido a transmitir:

- La parte que contiene información de la transmisión se entiende como la pública. En esta se establece, entre otra información, la IP que identifica el origen y destino de los paquetes de datos. La parte pública es analizable por ley (Telecommunications, n.d.).
- La parte que contiene el contenido de datos transferidos, es decir la conversación entre pares, se entiende como parte privada, cuya seguridad puede elevarse mediante encriptación.

Estas dos partes muestran un Internet en un primer estadio hacia la realidad distópica vivida hoy en cuanto a falta de privacidad digital (Goldberg, 2007; Goldberg, Wagner, & Brewer, 1997). Esta se entiende como la génesis de la infraestructura protocolaria de interconexión digital, que es de base abierta y de control. Cualquiera que esté escuchando puede saber quién envía datos a quién, con qué frecuencia, en qué momento e incluso el volumen de la conversación según cantidad de datos enviados. El objetivo es crear un perfil del navegante con el que conseguir rédito. A esta creación de perfiles se le llama *dossier effect* y los que mercadean con ellos *data brokers* (Rosenbaum, 1997). La importancia de los datos es notable y configura un estado de vigilancia activa (Zuboff, 2015) por los beneficios que aportan grandes cantidad de datos “perfilizados” (Mayer-Schönberger & Cukier, 2013).

Este primer estadio de interconexión digital de las personas, es decir, de la génesis del mundo virtual en el que las personas trascienden a cibernautas, no está diseñado por defecto para ser absolutamente privado (Seničar, Jerman-Blažič, & Klobučar, 2003). Si los cibernautas son aquellas personas que navegan por el ámbito digital,

ineludiblemente están desprovistos de privacidad al completo si no se usan procedimientos específicos para evitarlo, como Tor o FreeNet (Goldberg, 2007).

II.2.1.2. Web: HTTP y HTML

El segundo estadio de interconexión sucede con el surgimiento del protocolo *HyperText Transfer Protocol* (HTTP) (Berners-Lee, Fielding, & Frystyk, 1996) y el lenguaje de marcas *HyperText Markup Language* (HTML) (Berners-Lee, Connolly, Muldrow, & DTDs, 1986). El protocolo HTTP establece un modelo de comunicación entre dispositivos cliente y servidores de documentos. El lenguaje HTML facilita la creación de documentos digitales hiperconectados a partir de la información generada en la vida real.

En este segundo estadio se permite digitalizar la información disponible en papel, a modo de revolución evolucionada de la imprenta de Gutenberg. El binomio HTTP/HTML supone una disrupción digital. Permite trascender de lo estático real tangible a lo estático digital intangible. Los distintos documentos HTML hiperconectados crean la llamada *World Wide Web*, o simplemente *Web* (Berners-Lee, 1992). Esta revolución digital marca el camino hacia la interconectividad que vivimos actualmente.

El creador del HTTP y el HTML es Tim Berners Lee, que en sus propias palabras creó este sistema con la voluntad de ser abierto, afirma que la situación ha virado hacia un estado de poder centralizado en unos pocos y que debe revertirse. Así lo expresa en su nuevo proyecto Solid (Inrupt Inc., 2019):

Solid fue creado por el inventor de la World Wide Web, Sir Tim Berners-Lee. Su misión es remodelar la web tal y como la conocemos. Solid permite a los usuarios y organizaciones separar sus datos de las aplicaciones que los utilizan (Inrupt Inc., 2019).

A pesar de sus buenas intenciones, Berners Lee es, en gran medida, culpable de la situación al no proveer inicialmente a sus protocolos de medidas de seguridad o privacidad.

La definición del protocolo HTTP (Berners-Lee et al., 1996) sirve de oráculo para comprender en profundidad la desprotección de los cibernautas cuando comparten información de la vida real o la generan con la simple navegación en el ciberespacio. Así se explicita en el punto *12.3. Abuse of Server Log Information*:

Un servidor está en condiciones de guardar datos personales sobre las peticiones de un usuario que pueden identificar sus patrones de lectura o temas de interés. Esta información es claramente confidencial por naturaleza y su manejo puede estar restringido por la ley en ciertos países. Las personas que utilizan el protocolo HTTP para proporcionar datos son responsables de garantizar que dicho material no se distribuya sin el permiso de ninguna persona identificable por los resultados publicados (Berners-Lee et al., 1996).

En nuestro mundo digital interconectado los cibernautas generan trazas. Dichas trazas quedan almacenadas en registros dentro de los servidores web a las que acceden. Al utilizar el conjunto de las trazas almacenadas, aunque contengan datos personales anonimizados o seudonimizados, los cibernautas pueden ser identificados y desprovenerse de su privacidad (Weippl & Min Tjoa, 2005).

II.2.1.3. Plataformas

El tercer estadio se inicia con la web dinámica. Los avances en los *scripts* de navegador y de servidor convierten a cualquier página web en un canal de comunicación entre personas y bases de datos basada en formularios (Preibusch, Krol, & Beresford, 2013). Esto es posible por la evolución de los protocolos y tecnologías de la infraestructura Internet y Web. Los cibernautas pueden tanto leer información personal como enviarla a voluntad propia.

En este estadio nacen los foros, las redes sociales, las tiendas *online* y plataformas de bienes/servicios cuya información enviada por los cibernautas es consentida voluntariamente (Parker, Van Alstyne, & Choudary, 2016).

II.2.1.4. *Apps*

El cuarto estadio facilita el acceso abierto y continuado a plataformas. Tras la invención del *smartphone* (Egelman, Felt, & Wagner, 2013) se pone a disposición un acceso directo a la Web mediante aplicaciones llamadas *apps*. Las *apps* son una extensión de las plataformas y, por consiguiente, de la Web. Estas *apps* para móviles usan una serie de tecnologías web que conectan directamente con las plataformas. Incluso existen las llamadas *WebApps*, que son aplicaciones para móviles con un navegador incrustado con conexión directa a las plataformas web (Wetherall et al., 2011). En consecuencia, las *apps* permiten el acceso a un entorno digital de dudosa privacidad (Razaghpanah et al., 2018, 2015; Vallina-Rodriguez, 2017), en la que es muy fácil hacer un seguimiento sin que el usuario sea consciente de ello (Exodus Privacy, 2019).

II.2.2. *Clickstream*

Las páginas web son documentos HTML que permiten mostrar información hiperconectada. Berners-Lee (1992) imagina un contexto digital interconectado donde desde una página web se puede acceder a otra mediante palabras enlazadas. A estos enlaces les llama hiperenlaces y de esta manera un usuario navega por la web mediante clics de ratón encima de estos, generando un flujo de clics, en inglés *clickstream*.

El historial de navegación web consiste en los distintos accesos a las páginas web y se almacena en dos lugares:

- En el historial de navegación del navegador web del usuario.
- En los registros del servidor web en el que se almacena la página web visitada.

En consecuencia, los administradores de los registros del servidor web pueden reconstruir la navegación de los usuarios, descubrir identidades y recolectar información adicional sin que el usuario sea consciente (Berners-Lee et al., 1996). Cody lo expone muy claro de la siguiente manera:

Alimentar las preocupaciones de privacidad individual en línea es el hecho de que la recolección y el uso de información personal identificable nunca ha sido más barato o fácil que en el entorno en línea. Dicha información puede obtenerse de un usuario en línea de diversas maneras, con o sin el conocimiento del usuario. Por ejemplo, mediante el uso de una "cookie" o el seguimiento del "Clickstream" de un usuario, un sitio web (o "sitio") puede determinar la dirección de correo electrónico de un usuario, el tipo de ordenador que está utilizando, la información a la que el usuario accede en el sitio web y el tiempo que éste permanece en él, todo ello sin el consentimiento del usuario" (Cody, 1998).

Al transferir las anteriores afirmaciones de Cody (1998) en contexto educativo, se pone de manifiesto como la navegación web por clics, o *clickstream*, facilita la captura de datos personales de los estudiantes en entornos de aprendizaje en línea. *Clickstream* es el método fundamental en la aplicación de *Learning Analytics*. Se expone a continuación los distintos elementos coyunturales que influyen en el desarrollo de *Learning Analytics* y que son parte del problema de la investigación.

II.3. MOOC

Los MOOC son un contexto clave para el desarrollo del *Learning Analytics*. Analizar datos educativos se hace necesidad en la ejecución de los MOOC por el afán inicial de mejorar los rendimientos académicos.

En esta sección se expone el origen de los MOOC, su evolución hasta la actualidad y cómo el hecho de convertirse en banco de experimentación de datos educativos influye en el desarrollo del *Learning Analytics*.

II.3.1. Origen de los MOOC

Desde el primer curso MOOC conectivista (cMOOC) (Siemens, 2005, 2007) conducido por Siemens y Downes 2008 (2008), la evolución de estos cursos masivos ha llevado a la creación de todo un ecosistema de cursos impartidos por distintas entidades de un renombre considerable (Costa, Teixeira, & Alvelos, 2018). Hay razones de peso para

afirmar que la integración de los MOOC en educación superior ha provocado una revolución en su modelo de aprendizaje a distancia (Young, 2018; Yuan & Powell, 2013). Los primeros MOOC experimentales se dejan atrás, se consolidan como una auténtica disrupción educativa en modelos de aprendizaje a distancia y sigue innovándose su diseño con nuevas propuestas híbridas (García-Peñalvo et al., 2018). Esto se justifica por los avances en el modelo de negocio basado en MOOC, que muestra cómo las universidades estadounidenses se apresuran a ofrecer este tipo de cursos en línea gratuitos, con acreditaciones de pago e integrándolos en su programa de estudios. Este hecho arrastra a distintas universidades de todo el mundo a hacer lo mismo (Almatrafi & Johri, 2018).

La adopción de los MOOC en Estados Unidos se justifica por la situación crítica de los altos costes repercutidos en los estudiantes universitarios (Marcucci & Johnstone, 2007). Los estudiantes universitarios estadounidenses se han visto obligados a adquirir una deuda a largo plazo debido al alto costo de la matrícula y la compra obligatoria de libros de texto para acceder a los contenidos de las asignaturas (Epelboin, 2013). Esto ha provocado un descenso dramático en la tasa de estudiantes matriculados y, por tanto, se prevé una falta de profesionales en algunos sectores en la próxima década. De acuerdo con el informe de Indicadores 2012 de la Organización para la Cooperación y el Desarrollo Económico (OECD, 2012), solo el 42% de los jóvenes estadounidenses de 25 a 34 años tiene un título superior.

La aparición de los cMOOC propiciada por Siemens y Downes (2008) se ha considerado como el siguiente paso en la evolución de la educación abierta. Este tipo de cursos ha provocado una profunda reflexión en el sistema educativo tradicional, especialmente en lo que respecta a los modelos de docencia no presencial de las instituciones de educación superior (García-Peñalvo, 2016, 2019a). La educación en línea brinda oportunidades innovadoras a la educación superior, donde las universidades estadounidenses han logrado adaptar su modelo de negocio actual en un tiempo récord. Lo han hecho de acuerdo con sus propias necesidades y como ruta de escape a la crisis que enfrentan (Epelboin, 2013).

De esta adaptación emerge un nuevo tipo de MOOC, orientado a la adquisición de conocimientos académicos, más enfocado a la estandarización por evaluación por cuestionarios y al que se le distingue de los MOOC conectivistas como *Extension MOOC* (xMOOC). La base de los xMOOC es de aproximación epistemología conductual (J. Daniel, 2012). La Figura 3 muestra la evolución de los MOOC, desde la aparición de los cMOOC hasta su escisión en xMOOC.

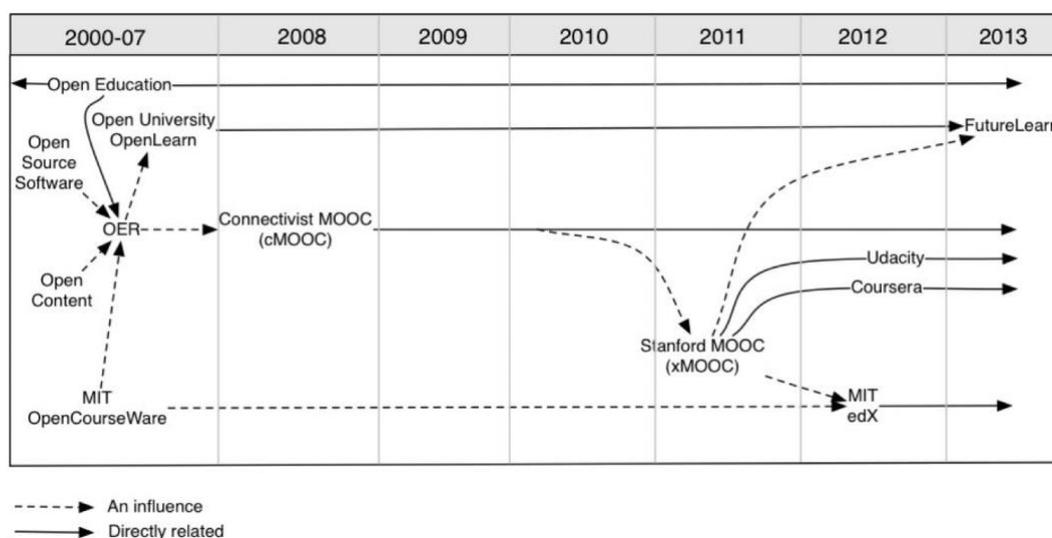


Figura 3 Línea de tiempo sobre MOOCs y Open Education. Fuente: (Yuan & Powell, 2013)

En relación con el término de "innovación disruptiva" acuñado por Christensen (Christensen & Bower, 1995; Clayton, 2015), los MOOC pueden considerarse innovaciones disruptivas. Estos permiten a toda una nueva población de consumidores, situada en la parte inferior de un mercado, acceder a un producto o servicio que históricamente solo era accesible para los consumidores con mucho dinero o habilidades de nivel superior.

Las universidades estadounidenses tienen un modelo de negocio sostenido por matrículas muy elevadas y contenido de aula distribuido en formato libro de papel, cuyos costes siguen la misma línea de encarecimiento. Estos servicios educativos se consideran caros e inaccesibles para ciertos estratos de la población. En consecuencia, los xMOOC rompen esta situación elitista y permiten a las universidades reconducir el contexto histórico y tradicional que prestan hasta hoy en día (Epelboin, 2013).

Las universidades también se ven obligadas a desagregar los elementos educativos como facultad, currículo y credenciales como nuevo modelo de negocios porque no pueden permitirse el cambio por sí solas. Finalmente, hay una razón convincente para creer que estos cambios disruptivos afectan a otros contextos distintos a las universidades estadounidenses. Esto es debido a la rápida afluencia y surgimiento de plataformas MOOC como MiríadaX (2014), UNIMOOC-AEmprende (Pedreño, Moreno, Ramón, & Pernías, 2013), MOOC UPC o MOOC UJI (2013). A pesar de la irrupción de estas plataformas, se deben realizar más investigaciones para asegurar que haya una accesibilidad real para toda la población de los países en desarrollo y no solo una parte de ella, como lo exponen Liyanagunawardena et al. (2013).

Los MOOC trazan un camino claro hacia la democratización de contenidos y acceso. Las universidades aprovechan dicha democratización y agregan un valor añadido a los cursos en línea gracias a su conocimiento. Por consiguiente, los MOOC se colocan en la última etapa de evolución de los recursos educativos abiertos (James Mazoue, 2013). Esto también hace que se democratizen los EVA y estos sean un elemento normalizado en el aprendizaje de todo estudiante, tanto en línea como presencial.

La apertura de contenidos es un beneficio que permite a los estudiantes acceder al conocimiento de profesores de alto reconocimiento de universidades como Stanford o Harvard. Además, el acceso abierto les permite a los estudiantes disfrutar de un proceso de aprendizaje único, pruebas auto evaluativas, actividades y relaciones heterogéneas en redes interconectadas con lo que consiguen un rico abanico de conocimiento y experiencia. Gracias al valor intrínseco de masividad de los MOOC, es posible acceder a un sistema de certificación a un costo muy por debajo en comparación con los altos precios de los grados universitarios.

Las universidades no son las únicas que han adoptado el modelo MOOC. Surgen distintas entidades privadas que ofrecen este tipo de cursos masivos. Entre estas se encuentran la ya citada Coursera, EdX o Udacity (Taneja & Goel, 2014). No obstante, y a pesar de su carácter independiente, todas ellas son originarias de universidades. Coursera y Udacity son fundadas por dos parejas distintas de profesores de la Universidad de Standford y EdX de la unión del MIT con la Universidad de Harvard.

A pesar de los desafíos a los que se enfrentan los MOOC en educación superior (Li, 2019), en estos momentos existen más de cien millones de personas cursando más de once mil MOOC ofrecidos por más de novecientas entidades educativas (Shah, 2019). La Figura 4 muestra el crecimiento de los MOOC desde el 2012.

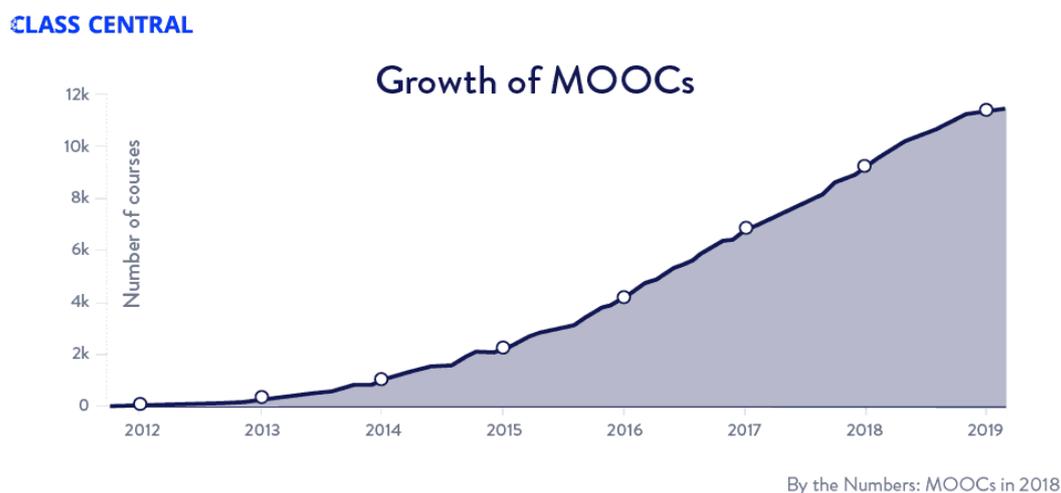


Figura 4 Crecimiento de los MOOCs desde el 2012. Fuente: (Shah, 2019)

II.3.2. Banco de datos para experimentación: minería de datos

La sección anterior muestra como de una experiencia entre dos profesores (Siemens & Downes, 2008) nace toda una cultura mundial de cursos masivos en línea que pretende revolucionar la educación superior. A propósito, se muestra todo su recorrido para indicar en mayúsculas que los MOOC son una gran fuente de datos para la comunidad de investigadores y empresas independientes. Asimismo, se destaca la sensibilidad aumentada del problema indicado en esta tesis, puesto que millones de trazas de estudiantes se generan constantemente y a tiempo real a lo largo de todos los MOOC activos y pasivos. Datos personales y generados en las plataformas MOOC están almacenados en servidores de distintas entidades, que acceden, transforman y comparten libremente.

Estos cursos son masivos, característica que los convierte en un medio de obtención de datos a tiempo real y de un altísimo valor. Por ejemplo, un curso sobre inteligencia artificial en la Universidad de Standford causó un registro de más de 160.000 estudiantes (Rodríguez, 2012). Las interacciones realizadas por estos alumnos proporcionan una gran cantidad de datos educativos de alto valor para su análisis y extracción de conclusiones para mejorar el aprendizaje (Romero & Ventura, 2007).

Anant Agarwal, presidente de edX en 2013 (Stokes, 2013), remarca y otorga un especial potencial a los MOOC para mejorar la pedagogía en entornos en línea. Al conjunto de datos recolectados de las escuelas de su consorcio y uso de técnicas de *Big Data* les llama “Partícula aceleradora del aprendizaje”. Esta clasificación es tendenciosa y demuestra la creencia de que el análisis de datos educativos son la base para la mejora del aprendizaje en entornos en línea. Demuestra que con los datos se puede gobernar el contexto educativo (Williamson, 2016a).

Del mismo modo que en entornos de negocio, los denominados *business intelligence* y *data analytics* (H. Chen, Chiang, & Storey, 2012) despiertan un fuerte interés alrededor del análisis de datos como mejora de procesos de negocio, en educación también surge un interés de análisis datos para la mejora del aprendizaje, e incluso con el uso de herramientas de negocio (Amo, Casañ, & Alier, 2014). Por consiguiente, se usa *data mining* para explotar las recolecciones de las interacciones de los estudiantes en todos los MOOC (Romero & Ventura, 2013). Este *data mining* en educación se proclama como *Educational Data Mining* (Romero & Ventura, 2007), del cual divergen dos principales vertientes analíticas como son *Learning Analytics* y *Academic Analytics* (ver II.4 *Educational Data Mining*).

Usar el análisis de datos educativos permite al equipo detrás de un MOOC dar respuesta a las siguientes preguntas y, posiblemente, corregir la situación para evitar el abandono, los bajos resultados y ofrecer un entorno de aprendizaje adaptativo (Clark, 2013; Romero & Ventura, 2010): ¿cómo diferentes estudiantes eligen usar diferentes recursos de aprendizaje y obtener diferentes resultados?, ¿la clase entiende el material lo suficientemente bien como para continuar?, ¿alguno de los estudiantes requiere instrucción de recuperación?, o ¿qué estudiantes probablemente necesitarán

asesoramiento académico para completar la escuela con éxito? Algunos autores dan respuesta a las preguntas anteriores con el análisis de las trazas generadas en los MOOC (Kizilcec, Piech, & Schneider, 2013; Tabaa & Medouri, 2013; Yang, Sinha, Adamson, & Rose, 2013).

II.3.3. Mejora de rendimientos con *Clickstream*

Clickstream (ver sección II.2.2 *Clickstream*) es uno de los métodos de recolección usados por autores que en sus investigaciones realizan análisis de datos educativos. Este proceso es debido a la arquitectura web en la que se conducen los MOOC y a su facilidad de captura de información del comportamiento a través de los clics en los hiperenlaces. En un inicio se usa el *clickstream* para la mejora de los resultados académicos (Brinton, Buccapatnam, Chiang, & Poor, 2016; Brinton & Chiang, 2015; Yu et al., 2019). Incluso se usa para perfilar a los estudiantes y generar visualizaciones para ayudar al profesorado a mejorar el seguimiento, tutoría y evaluación de los estudiantes como se muestra en la Figura 5 (Amo, Casañ, et al., 2014).

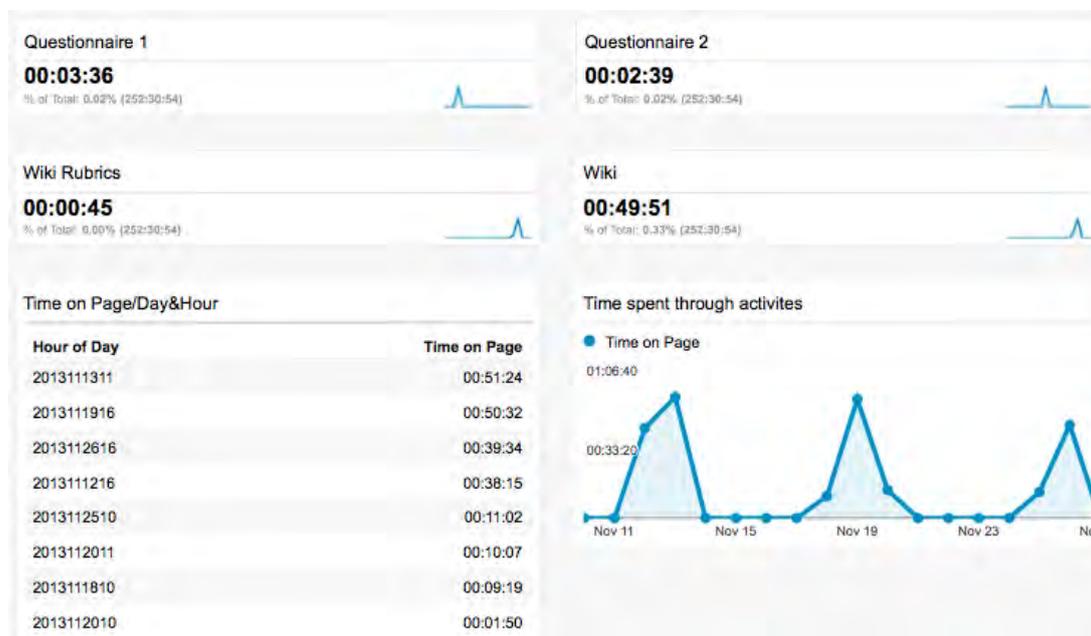


Figura 5 Dashboard de un estudiante en base al tiempo de dedicación. Fuente: (Amo, Casañ, et al., 2014)

Afortunadamente autores como Guàrdia et al. (2013), Mazoue (2013), de Waard (2013), Glance, Forsey y Riley (2013), Sonwalkar (2013), entre muchos otros, han dedicado sus esfuerzos a establecer fundamentos de diseño y pedagogía de los MOOC, con el fin de

cumplir con los criterios de calidad para convertirlos en un entorno de aprendizaje adaptable y eficaz, más allá de potenciar el rendimiento académico de los estudiantes. No obstante, ninguno acude a las leyes de protección de datos como reglamento base del que partir. No es hasta el 2014 que en *Learning Analytics* empiezan a surgir preocupaciones éticas y de privacidad (Pardo & Siemens, 2014).

Antes de exponer cuestiones sensibles sobre la privacidad de los datos de los estudiantes en el uso de *Learning Analytics* se procede a exponer el campo científico al que pertenece, en qué consiste y con qué otros acercamientos de análisis educativos se complementan.

II.4. Educational Data Mining

Learning Analytics se entiende como una rama de *Educational Data Mining* (Romero & Ventura, 2010), incluso algunos los equiparan conceptualmente (Berland, Baker, & Blikstein, 2014), y para otros tiene una parte de *Educational Data Mining*. La similitud proviene de los métodos analíticos, donde *Educational Data Mining* es anterior a *Learning Analytics* y esta los adopta *a posteriori* como recurso educativo. A pesar de las similitudes metodológicas, ambas aproximaciones tienen objetivos educativos distintos, aspecto clave que las diferencia.

Romero (Romero & Ventura, 2007) describe el objetivo de *Educational Data Mining* como:

...utilizar conjuntos de datos educativos a gran escala para comprender mejor el aprendizaje y proporcionar información sobre el proceso de aprendizaje... (Romero & Ventura, 2013).

Siemens et al. (2011) describen el objetivo de *Learning Analytics* como:

...la medición, la recopilación, el análisis y la presentación de informes de los datos sobre los educandos y sus contextos, con el fin de comprender y optimizar el aprendizaje y los entornos en los

que éste tiene lugar. Los análisis de aprendizaje se centran en gran medida en mejorar el éxito de los estudiantes... (Siemens et al., 2011).

La diferencia entre *Learning Analytics* y *Educational Data Mining* adoptada por distintos autores es el objetivo de aplicación. Duval (2012) lo expone en su creencia de que la diferencia radica en el aprendizaje en lugar de la computación algorítmica:

... *Learning Analytics* consiste en recoger las trazas que los estudiantes dejan atrás y utilizarlas para mejorar el aprendizaje. *El Educational Data Mining* puede procesar las trazas algorítmicamente y señalar patrones o calcular indicadores. Mi interés personal es más bien utilizar las trazas y capacitar a los estudiantes para que sean "mejores estudiantes" (Duval, 2012).

De aquí que el trabajo de Duval se fundamente en la visualización de datos para ayudar a la mejora del aprendizaje (Charleer, Klerkx, Duval, De Laet, & Verbert, 2016).

II.4.1. Learning Analytics

A pesar de algunas definiciones adoptadas por defecto, en realidad el concepto de *Learning Analytics* aún está en proceso de definición. La primera referencia a *Learning Analytics* aparece en Wikipedia el 23 de agosto del 2010 ("Learning Analytics," 2010). Define *Learning Analytics* como "el uso de datos y modelos para predecir el progreso y rendimiento de los estudiantes y la habilidad de actuar con esa información". Esta definición tiene una influencia directa del *Educational Data Mining*, y en realidad surge de la página web de nextgenlearning.com (Next Generation, 2010), datada en 11 de Julio del 2010. La entrada creada en Wikipedia es muy relevante, puesto que es resultado del grupo de discusión que abre George Siemens para debatir sobre este nuevo concepto de análisis de datos educativos ("Learning Analytics Google Groups," 2010).

Siemens, que aboga por el conectivismo (Siemens, 2005), es el primero en entrar en debate académico para definir *Learning Analytics* desde una perspectiva pedagógica

(Siemens, 2010c). A las definiciones de entonces les añade datos más allá del EVA, como pueden ser redes sociales o blogs personales. Describe la analítica del aprendizaje como:

El uso de datos inteligentes, datos producidos por los estudiantes y modelos de análisis para descubrir información y conexiones sociales, y para predecir y aconsejar sobre el aprendizaje (Siemens, 2010c).

La definición de Siemens es más amplia en el sentido de que se adapta al sistema educativo y lo complementa. Su descripción, en sus propias palabras, “es menos limpia, pero no intenta modificar el sistema educativo”, sino utilizar los resultados analíticos para mejorarlo.

La definición de *Next Generation* se orienta al uso exclusivo de la minería de datos. En cambio, la definición de Siemens es más conectivista y orientada a ampliar el espectro de *Learning Analytics*. Como la amplía más allá de considerar aspectos analíticos, Siemens indica que *Learning Analytics* concierne a otros conceptos como acciones, definición del currículum, adaptación y personalización o predicción (ver Figura 6).

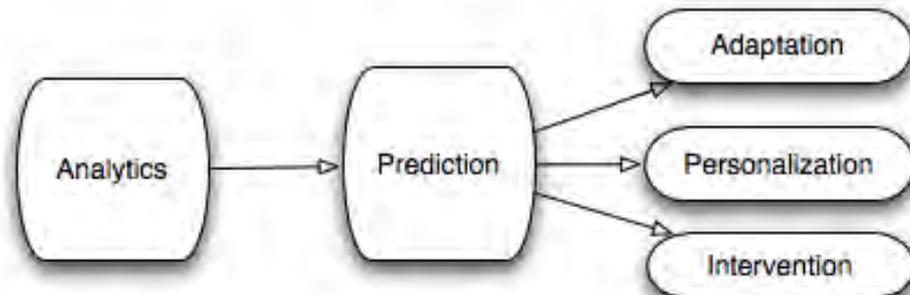


Figura 6 Objetivos de uso de los datos en Learning Analytics. Fuente: (Siemens, 2010c)

Se nota aún el predominio del campo de minería de datos en el sentido que Siemens explica *Learning Analytics* como un proceso puramente predictivo.

Siemens establece el 2010 como el año en el que el concepto *Learning Analytics* toma presencia en el mundo educativo. Al mismo tiempo se apunta a Siemens como el máximo impulsor del término, tras su trayectoria de cursos conducidos alrededor de

esta temática (Siemens, 2010b), de congresos *Learning Analytics and Knowledge* y charlas, pero sobre todo por la constitución de *Society for Learning Analytics Research* (SoLAR) (SoLAR, 2019), en el que fue Presidente Fundador entre el 2013 y el 2015, y donde ejerce actualmente de miembro activo.

Según Chatti et al. (2012), el objetivo de *Learning Analytics* es, principalmente, mejorar y optimizar el contexto educativo a partir de conceptos y métodos provenientes de distintos campos. Para entender mejor el concepto, Chatti et al. muestran un proceso cíclico de tres pasos distintos (ver Figura 7):

1. Recopilación y pretratamiento de datos: Representa la recopilación de todos los datos necesarios para un método en particular y su transformación a un formato adecuado para ese método.
2. Análisis y acción: Incluye el análisis y visualización de la información extraída del primer paso y las acciones sobre esa información.
3. Post-procesamiento: Basado en la mejora continua, el post-procesamiento incluye todos aquellos cambios en el método basados en los dos pasos anteriores.



Figura 7 Proceso de Learning Analytics. Fuente: (Chatti et al., 2012)

En resumen, y como primer estadio, se puede afirmar que *Learning Analytics* es una estrategia cuantitativa que analiza las interacciones de los estudiantes en EVA. En un segundo estadio se empiezan a analizar los estudiantes en el aula física o en el uso de los dispositivos como complemento al análisis de EVA.

La rama del *Learning Analytics* que trasciende el análisis de las plataformas a la realidad se llama *Multimodal Learning Analytics* (Blikstein, 2013; Ochoa et al., 2017) y consiste en comprender cómo los estudiantes interactúan con dispositivos físicos mediante gestos, posturas, movimientos e incluso expresiones faciales. Algunos de los ámbitos de estudio de esta rama pueden verse en la Tabla 1 y un ejemplo de dispositivo de captura de datos en el aula en la Figura 8.

Tabla 1 Contextos de estudio de *Multimodal Learning Analytics*. Fuente: Extraído de (Ochoa et al., 2017)

Contexto	Modo
Lectura	Movimiento, mirada, gestos, expresión facial, habla
Presentaciones orales	Postura, movimiento, gestos, mirada, habla, documento digital
Resolución de problemas	Movimiento, acciones, hablar, escribir, dibujar
Ejercicios constructivos	Gestos, acciones, habla, expresiones faciales, respuesta cutánea galvánica
Uso de tutores inteligentes	Archivos de registro digitales, expresiones faciales, voz



Figura 8 Diseño de un dispositivo de grabación multimodal para ser utilizado en entornos de lectura: en la sala de clase (izquierda) y desde el punto de vista del estudiante (derecha). Fuente: (Ochoa et al., 2017)

El *Multimodal Learning Analytics* se vuelve más intrusivo con elementos de detección facial (You, 2019) o incluso de captura de ondas cerebrales (Jing, 2019). En este nuevo estadio de comprensión del aprendizaje de los estudiantes se recolectan una serie de datos biométricos que los identifican sin lugar a duda, hecho que explica una mayor sensibilidad a medida que las tecnologías avanzan y se integran en el ámbito educativo.

II.4.2. *Academic Analytics*

Sin ser de ámbito de estudio en la presente investigación se considera citar brevemente la línea de mejora educativa entendida como *Academic Analytics*, el cual consiste en:

...la mejora de los procesos organizativos, los flujos de trabajo, la asignación de recursos y la medición institucional mediante el uso de datos de los estudiantes, académicos e institucionales. La analítica académica, similar a la analítica de negocios, se preocupa por mejorar la efectividad organizacional (Siemens et al., 2011).

Las diferencias entre *Academic Analytics* y *Learning Analytics* se muestran en la Tabla 2. La primera aproximación analítica tiene un nivel institucional con alcance hacia arriba institución-gobierno y la segunda aproximación analítica tiene un nivel departamental con alcance hacia abajo departamento-aula.

Tabla 2 Diferencias entre *Academic Analytics* y *Learning Analytics*. Fuente: (Long & Siemens, 2011)

Tipo de analítica	Nivel u objeto de análisis	¿Quién se beneficia?
<i>Learning Analytics</i>	Nivel de curso: redes sociales, desarrollo conceptual, análisis del discurso, "currículum inteligente"	Estudiantes, Profesores
	Departamental: modelado predictivo, patrones de éxito/fracaso	Estudiantes, Profesores
<i>Academic Analytics</i>	Institucional: perfiles de estudiantes, rendimiento académico, flujo de conocimientos	Administradores, Patrocinadores, <i>Marketing</i>
	Regional (estado/provincial): comparación entre sistemas	Patrocinadores, <i>Marketing</i>
	Nacional e Internacional	Gobiernos locales, autoridades educativas

II.4.3. Social Network Analytics

Desde que la Web 2.0 (O'Reilly, 2007) y las tecnologías sociales se integran en el sistema educativo, se realizan muchos esfuerzos de investigación para comprender interacciones entre estudiantes y mejorar la educación en línea. La evolución de los EVA hacia entornos más sociales, como los introducidos por los MOOC (Baggaley, 2013; Haggard, 2013), es tendencia por los beneficios de colaboración social entre pares (Gracia, Gil, & Osinaga, 2012).

Para comprender las interacciones educativas en tecnologías sociales y colaborativas se requieren nuevas metodologías de análisis. Por consiguiente, aparecen nuevos enfoques de *Learning Analytics* o *Educational Data Mining* (Harmelen & Workman, 2012). Por un lado, los profesores son capaces de comprender y optimizar los procesos de aprendizaje dentro de sus cursos. Por otro lado, los estudiantes disponen de un tutor aumentado y un aprendizaje más personalizado (Bienkowski, Feng, & Means, 2014). Por tanto, para analizar los entornos de aprendizaje social se necesita un método de análisis social (Adkins, 2009). *El Social Network Analytics* (SNA, en castellano Analítica de Redes Sociales) es considerado una rama del *Learning Analytics* y la principal metodología para lograr tal cometido (Siemens, 2005).

Se realizan esfuerzos para integrar análisis sociales en EVA como Moodle. En este sentido se desarrollan herramientas con propósitos educativos, tanto internas (Chan, 2014) como externas (Bakharia, Heathcote, & Dawson, 2009), e incluso se utilizan herramientas no meramente educativas como Gephi (Bastian, Heymann, & Jacomy, 2009). Feldstein (2010) nota una demanda real para el integrando de este tipo de herramientas de análisis social como una funcionalidad por defecto.

El tratamiento de los datos de los estudiantes en métodos de SNA también permite identificarlos y perfilarlos. El SNA dibuja un contexto mucho más sensible en cuanto al tratamiento de datos de los estudiantes, puesto que permite identificarlos y perfilarlos con inferencias relacionadas con el tono y sentimiento de las conversaciones. A pesar de las posibilidades educativas del SNA, ver Tabla 3, la información resultante del análisis implica una mayor sensibilidad en cuanto a su tratamiento y almacenado.

Tabla 3 Utilidades de Social Network Analytics. Fuente: (Amo, García-Peñalvo, & Alier, 2014)

Utilidad	Descripción
Identificar Patrones de Interacción o Relaciones entre los estudiantes	El SNA permite estudiar, describir e inferir las interacciones entre individuos en una red social, tanto a nivel individual como considerando un grupo o grupos de actores. De esta manera, como hay más información sobre las relaciones entre los estudiantes en la red de aprendizaje social de cualquier curso, el profesor puede tomar medidas en caso de que surja algún problema, como, por ejemplo, una posible deserción escolar (Badge, Johnson, Moseley, & Cann, 2011; Dawson, 2008; Tirado, Aguaded, & Hernando, 2011)
Identificar o mejorar las características de los estudiantes	El estudio de una red social puede detectar situaciones tanto positivas como negativas, pero también habilidades o características de los estudiantes que podrían ser controladas y mejoradas (Capuano, Mangione, Mazzoni, Miranda, & Orcioli, 2014; Shea et al., 2013)
Optimizar los entornos de aprendizaje social	Uno de los objetivos del SNA en educación es optimizar los roles educativos y personalizar los procesos de aprendizaje de los estudiantes. Ayuda a descubrir mejores formas de hacerlo (Erlin, Yusof, & Rahman, 2008; Hamre & Vidgen, 2008; Morueta, Gómez, & Gómez, 2011; Penuel, Korbak, & Hoadley, 2006)

A lo largo del capítulo se constata un claro tratamiento de datos educativos mediante su recolección automatizada, con afán de mejora y optimizado del contexto de enseñanza y aprendizaje, o comprensión de este. En conjunto, todas las aproximaciones anteriores dibujan un contexto sensible ante la captura y tratamiento de datos educativos. Surgen algunas preocupaciones éticas, de privacidad y confidencialidad de

datos, legales, técnicas y de seguridad con respecto a los datos recopilados. Estos problemas aumentan la desconfianza general en el uso del *Learning Analytics*.

II.4.4. Miedos y recelos: una cuestión delicada

Learning Analytics implica el tratamiento de datos sobre el comportamiento de los estudiantes, que pueden ser, incluso, menores de edad. Por tanto, se trata de una cuestión muy delicada. El ciclo analítico implica recolectar datos, almacenarlos durante largas temporadas y utilizarlos para realizar análisis y visualizaciones. Tales análisis pueden ser descriptivos, predictivos e incluso prescriptivos, hecho que supone gestionar, tratar y usar datos de carácter personal. Este contexto es de una alta sensibilidad a diferencia de los contextos individuales en los que se utiliza la analítica a voluntad propia.

Existe una creciente preocupación por el impacto analítico que tiene en los individuos, su identidad e integridad (Herold, 2014; Singer, 2014). Un ejemplo de estas preocupaciones se puede encontrar en el cierre de inBloom (Herold, 2014), debido a los problemas de privacidad planteados por los padres y los grupos de presión sobre sus prácticas de *Learning Analytics* y *Big Data* en la educación (Laveti, Kuppili, Ch, Pal, & Babu, 2017; Singer, 2014). En consecuencia, muchas situaciones se adhieren a marcos y políticas propuestos por investigadores para generar soluciones ad-hoc en las propias instituciones educativas (Tsai, Moreno-Marcos, Tammets, Kollom, & Gašević, 2018).

Existe un debate académico en curso sobre los enfoques que deberían adoptarse para abordar cuestiones acerca de la aplicación de *Learning Analytics* con respecto a la ética, la privacidad (Pardo & Siemens, 2014), los marcos jurídicos y los aspectos técnicos. Según Drachler y Greller (2012) hay temores en cuanto al *Learning Analytics*. Estos temores tienen sus raíces en:

- Cuestiones de privacidad e identidad digital.
- Relación de poder asimétrica que conlleva entre el responsable del tratamiento y el interesado.
- Propiedad de los datos sobre los estudiantes.
- Integridad de los datos.

- Seguridad de datos y riesgo de *hacking*.
- Problema del anonimato.

Actualmente, no hay soluciones técnicas definitivas para evitar todos estos problemas. Iniciativas de estandarización como IMS Caliper (Sakurai, 2014) o xAPI (Del Blanco, Serrano, Freire, Martinez-Ortiz, & Fernandez-Manjon, 2013) han abordado los problemas relacionados con la interoperabilidad de los datos, pero no abordan cuestiones como el miedo hacia el análisis del aprendizaje, arraigado en factores humanos, como la angustia, el escepticismo, los malentendidos y las preocupaciones críticas (Pardo & Siemens, 2014).

La lista de control DELICATE propuesta por Drachsler como resultado de un proyecto de investigación internacional (Drachsler, 2016) introduce ocho puntos por los que se debe pasar antes de iniciar un proyecto o una actuación en *Learning Analytics*:

- Determinación: Decidir el propósito de *Learning Analytics* para su institución.
- Explica: Definir el alcance de la recolección y uso de los datos.
- Legítimo: Explicar cómo opera dentro de los marcos legales, haciendo referencia a la legislación esencial.
- Involucrar: Hablar con las partes interesadas y asegurarles la distribución y el uso de los datos.
- Consentimiento: Buscar el consentimiento a través de preguntas de consentimiento claras.
- Anonimizar: “Des-identificar” a las personas en la medida de lo posible.
- Aspectos técnicos: Controlar quién tiene acceso a los datos, especialmente en áreas con alta rotación de personal.
- Socios externos: Asegúrese de que las aplicaciones externas proporcionen los más altos estándares de seguridad de datos.

La lista de control DELICATE muestra la complejidad del problema. Cada punto plantea una serie de preguntas difíciles y las respuestas pueden no satisfacer a todos los actores o partes interesadas. En el mejor de los casos, cuando todas las partes interesadas llegan a un entendimiento sobre los ocho puntos de la lista de verificación DELICATE, no hay

soluciones técnicas definitivas que hagan cumplir los acuerdos, así el error humano o los malentendidos pueden conducir al punto de partida.

SHEILA (Tsai et al., 2018) es otro proyecto destinado a ayudar a integrar *Learning Analytics* en instituciones educativas. Proponen un marco de desarrollo de políticas que junto a las instituciones permita integrar la analítica del aprendizaje de forma segura. LALA Community (Maldonado & Hilliger, 2018) es un claro ejemplo de transferencia de conocimiento del proyecto SHEILA en países latinoamericanos. Otras iniciativas como JISC (Sclater & Biley, 2015), una organización sin ánimo de lucro del Reino Unido, proponen un código de práctica para *Learning Analytics*.

Es por lo anterior que en la presente investigación se realizan los pasos necesarios para comprender cómo se percibe el problema por parte de autores y agentes de referencia, cuál es el foco de interés en el campo de la privacidad en procesos de *Learning Analytics* y qué soluciones existen en el campo científico para solucionar los problemas de desconfianza.

II.5. *Learning Analytics* y Conocimiento en congresos

Learning Analytics ha tenido un gran impacto e interés en la comunidad científica desde su introducción en el 2010 hasta la actualidad (Joksimović, Kovanović, & Dawson, 2019) (ver Figura 9).

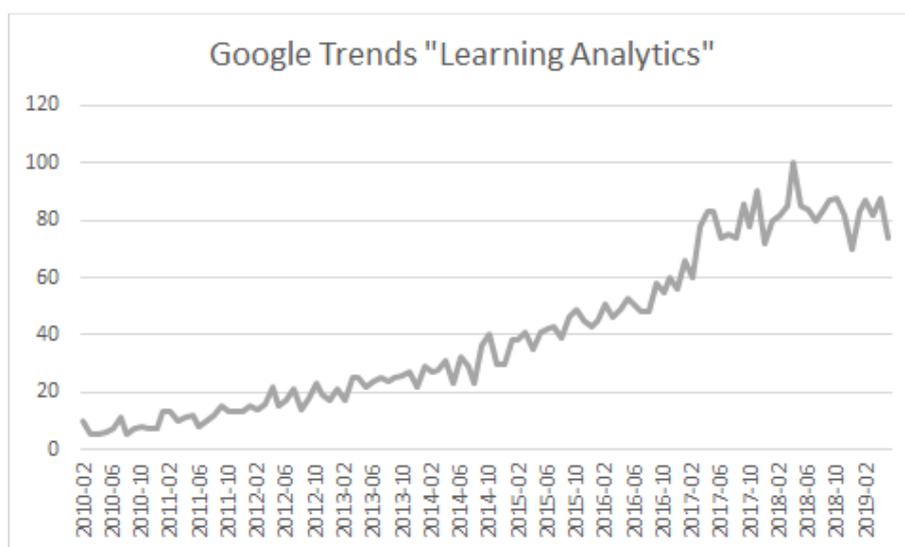


Figura 9 *Learning Analytics* en Google Trends, 01 de junio del 2019.

Este interés ha ido acompañado de un gran volumen de investigaciones, cuyos autores han intentado definir sus usos y sus distintas definiciones. A pesar de los esfuerzos y varias definiciones aceptadas por defecto, aún no hay un consenso total y definitivo sobre qué consiste exactamente el concepto, como se ha visto previamente. Por consiguiente, *Learning Analytics* aún está en proceso de madurez, incluso cuestiones pedagógicas (Matcha, Ahmad Uzir, Gasevic, & Pardo, 2019).

Adoptar ampliamente la práctica de *Learning Analytics* en educación supone superar una serie de retos, como la adopción de la competencia digital y la competencia para el tratamiento de datos (*data literacy*) por parte de los docentes (Mandinach & Gummer, 2013). Este puede ser el motivo por el que en los últimos 10 años la investigación en *Learning Analytics* ha crecido, pero no así su transferencia práctica en educación, cuyo impacto está siendo mucho más lento (Aznarte, Hidalgo, Rubió, & Ruipérez, 2019).

La preocupación por la confidencialidad, seguridad y buen uso de los datos educativos recolectados no suscitó interés hasta pocos años después de su aparición (Pardo & Siemens, 2014). Algunos autores entienden que el uso de datos de estudiantes es un asunto sensible que requiere de una profunda y cuidada atención desde distintas disciplinas. Aspectos como la ética, privacidad o seguridad de los datos entran en debate provocando que distintos autores empiecen a investigar a su alrededor (Drachsler, 2016; Hoel & Chen, 2016).

Una gran preocupación recae en la adecuada confidencialidad y seguridad de datos educativos de los estudiantes, que también pueden ser menores de edad. Es una cuestión importante y sensible que requiere de un acercamiento en profundidad. *Learning Analytics* precisa la recolección de datos de las interacciones de los estudiantes para después analizarlos con el objeto de extraer conocimiento, en inglés *insights* (Romero & Ventura, 2013). En algunos casos se están manipulando datos de menores, hecho que supone una sensibilidad adicional. Los últimos sucesos en cuanto a malos usos de datos personales (Herold, 2014) y las publicaciones de nuevos reglamentos por parte de autoridades gubernamentales (EP and the CEU, 2016) ponen de manifiesto la necesidad de regular por ley el uso de datos educativos en herramientas de *Learning Analytics*. Por tanto, es fundamental asegurar la privacidad y seguridad dichos procesos.

Es interesante conocer las tendencias en torno a la privacidad y seguridad publicadas en los congresos *Learning Analytics and Knowledge*, puesto que ahí se concentran los máximo dirigentes de este movimiento analítico. Conocer las conversaciones y debate en estos congresos alrededor de la temática “privacidad y seguridad en el tratamiento y uso datos educativos en *Learning Analytics*” aportan conocimiento para comprender el estado de la cuestión.

II.5.1. Congresos Learning Analytics and Knowledge

Los congresos *Learning Analytics and Knowledge* tienen sus orígenes en el año 2011 (Siemens, 2010a). Estos congresos tienen una alta relevancia, puesto que ahí se concentran los máximo dirigentes de este movimiento analítico, se exponen las distintas tendencias de uso del *Learning Analytics*, y se origina la discusión sobre su práctica y validez.

En cada encuentro se presentan investigaciones que ponen de manifiesto en qué ámbitos hay interés de usar el análisis de datos educativos. En el contexto de la presente investigación y el ámbito de la privacidad, se realiza una revisión analítica de los artículos presentados a lo largo de los congresos LAK. Es interesante conocer las conversaciones y debate en estos congresos alrededor de la temática “privacidad y seguridad en el tratamiento y uso datos educativos en *Learning Analytics*”. De los resultados se explica el objetivo de la investigación “comprender la necesidad de utilizar una tecnología protectora de la privacidad e identidad de los estudiantes”.

II.5.1.1. Revisión de la literatura LAK

A modo de comprensión de la literatura en congresos LAK, se realiza una revisión desde una posición analítico-textual de los títulos y resúmenes de las publicaciones. Para realizar la revisión se recolectan todos los títulos de los artículos presentados a lo largo de los congresos LAK celebrados entre 2001 y 2019. Las Actas de cada uno de los congresos están publicados en la ACM Digital Library (Long, Siemens, Gráinne, & Gašević, 2011). Mediante un proceso de *data scraping* se descargan los títulos, los

resúmenes y los autores de cada uno de los títulos. Se usan expresiones regulares para extraer los datos y desarrollar los algoritmos analíticos.

Una vez recolectados los datos de los artículos, se desarrollan una serie de algoritmos en Python para ejecutar un análisis textual de los títulos y resúmenes. Usando la librería *Natural Language Toolkit* (NLTK Project, 2019). El análisis de los títulos, del que se extraen términos más citados, ofrece la oportunidad de conocer qué aspectos son relevantes en cada uno de los LAK.

- Expresión regular para extraer los títulos: `(?:<td .*">)(.*?)(?:</a.*td>`
- Expresión regular para extraer los autores: `(?:<a .*author.*">)(.*?)(?:`
- Expresión regular para extraer los resúmenes: `(?:<span .*<p>)(.*?)(?:</p>`

Los resultados de los análisis mostrados en la Tabla 4 (se usa “...” cuando el número de términos es excesivo y su muestro es despreciable para las conclusiones de la investigación) muestran que los términos relacionados con “*privacy*” se detectan solo en un 1,13% (7 de 620) de los títulos publicados, cuya aparición va desde el 2015 al 2019. Los términos relacionados con “*ethic*” se detectan solo en un 0,97% (6 de 620) de los títulos publicados, cuya aparición va desde el 2012 al 2018. Por consiguiente, no existe un fuerte interés en aspectos éticos o de privacidad en los congresos LAK. La Figura 10 muestra un resumen de los artículos encontrados para los términos relacionados con “*privacy*”, “*ethic*” y “*security*”.

2011-19	
620	artículos publicados
6	ethic
7	privacy
1	security

Figura 10 Resumen de términos y artículos en los congresos LAK. Elaboración: propia.

A pesar de estos bajos resultados, y mediante la extracción y análisis de los resúmenes, se afirma que hay una clara preocupación desde el equipo directivo, puesto que los que se pronuncian en los artículos son autores de peso dentro de la comunidad LAK como son Siemens, Pardo o Drachsler (Drachsler & Greller, 2016; Lang, Siemens, Wise, &

Gasevic, 2017). No obstante, las palabras de los autores se quedan en teorías más que en hechos.

Tabla 4 Resultado de los análisis de texto de los artículos LAK11-19

Congreso	Términos
LAK11	(2) systems, predictive, online, education, data, case, study, exploratory, towards, dynamic, students, applying (1) learnometrics, metrics, objects, attention, visualization, recommendation, networks, ethical...
LAK12	(9) data (8) online, student, using (5) educational (4) building, model, tool, success (3) communities, meets, mining, workplace, open, case, study, course
LAK13	(7) assessment, data (5) case (4) online, study (3) evaluation, social, collaborative, using, analyzing, open (2) addressing, learner, issues, topic, models, discussions, pedagogical, intervention, student, multimodal, workshop... (1) ethical...
LAK14	(10) student (6) using (5) course, patterns, system (4) knowledge, study, performance, education (3) assessment, educational, success, data, visualizing, early, students, social, analyzing, discourse (1) ethical
LAK15	(12) student (9) performance, study (8) students, approach (7) data, using (2) privacy (1) ethical
LAK16	(16) data (15) student (14) using (9) students (8) case, MOOC, performance (2) privacy (1) ethical
LAK17	(20) using (18) student (16) data (10) students (9) MOOC (2) privacy
LAK18	(11) students (6) MOOC, using (4) practice, design, frameworks, writing, student, multimodal (3) knowledge, scale, educational, environment, behavior, test, dashboard, study, evaluation, online, research, strategies, data... (2) evaluating, retention, MOOC, assessment, teachers, model, feedback, tool, method... (1) ethics
LAK19	(13) student (6) using, students, knowledge (5) data, reading, models, online, feedback (4) predicting, exploring, predictive, engagement, model,

impact (3) time, study, course, facilitate, deep, trajectories, comprehension, MOOC, design... (1) privacy

Términos relacionados con “*security*” no aparecen mencionados en ninguno de los 620 títulos de los congresos LAK11-19. Solo aparece la palabra “*security*” en el resumen de un artículo del LAK18 y cuyo foco es el uso de la tecnología *blockchain* (Ocheja, Flanagan, & Ogata, 2018). Se destaca que la comunidad científica, incluso artículos publicados por los autores de esta investigación (Amo, Fonseca, Alier, García-Peñalvo, & Casañ, 2019), apunta a *blockchain* como una tecnología emergente que no asegura una adecuada privacidad y seguridad de los datos sensibles de estudiantes. Por consiguiente, se llega a la conclusión de que la seguridad tampoco es aspecto de interés en los investigadores LAK para resolver los problemas de privacidad de los que adolece *Learning Analytics*.

La Figura 11 muestra un *small multiples* de la nube de palabras generada a partir de los términos más citados en títulos y resúmenes de los artículos publicados en los congresos LAK desde la edición de 2011 a la edición de 2019.

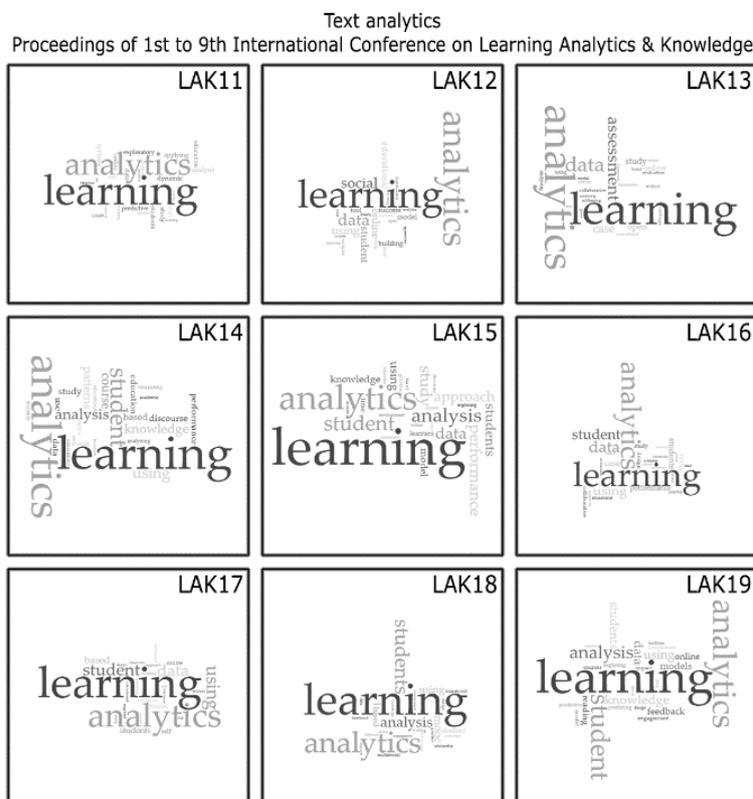


Figura 11 *Small multiples* de los términos más repetidos en los títulos LAK11-19. Fuente propia.

II.6. Leyes sobre protección de datos personales

Se requiere conocer las leyes que protegen los datos personales de los estudiantes, cuál es su percepción en los roles educativos y cuál es el nivel de adopción en instituciones educativas para poder proporcionar una solución válida a la problemática en discusión.

Los datos personales de los estudiantes tratados en procesos de *Learning Analytics* están regidos por distintas leyes de protección de datos. En este capítulo se presentan distintas leyes de protección de datos personales en el marco europeo y de España, a las que se hará referencia para exponer conclusiones futuras.

En España, las distintas leyes que orbitan alrededor de los datos recolectados por cualquier entidad, incluso no educativa, tienen una historia de más de 20 años llena de modificaciones y derogaciones (Sánchez Bravo, 1994). En el marco de la educación, el contexto legal ayuda a superar, a alimentar y a entender la desconfianza existente alrededor del tratamiento de los datos educativos recolectados (Singer, 2014). Esta desconfianza existe más allá del país donde discurren los procesos de aprendizaje. Aun estableciendo vínculos de protección entre países, la desconfianza traspasa fronteras y hay quien teniendo asegurado el marco legal evita cualquier transferencia de datos y, por tanto, uso de herramientas tecnológicas con usos educativos (XNET, 2019).

Esta situación de desconfianza es debida en parte al desconocimiento total o parcial del contexto legal, a las posibles ambigüedades y a las distintas derogaciones o actualizaciones de las leyes. El lenguaje jurídico tampoco ayuda a entender las leyes para quienes no están familiarizados, este es el caso de las regulaciones sobre las *cookies* y los problemas de privacidad derivados (Leenes & Kosta, 2015; Zimmerman, 2001).

Un contexto legal de por sí, aún acompañado de medidas disuasorias, no es una herramienta eficaz contra el mal uso o captación de datos personales de los estudiantes -al menos hasta que no haya denuncia- (Robinson, 2017). Esta coyuntura aumenta el estado de desconfianza ante el uso de *Learning Analytics*.

Se suma un aumento de desconfianza en aquellas herramientas que prometen unas funcionalidades a coste cero, que usan y tratan datos de alumnos (XNET, 2019). Ya han

aparecido casos de escuelas tecnológicas que están en el punto de mira por considerar modelos de negocio que comercian con datos de alumnos, incluso algunos han cerrado (Robinson, 2017). No obstante, sea el producto o precio que ofrezca cualquier empresa que recolecte datos personales de alumnos debe aplicar las leyes de su país y de los que capta datos.

A continuación, se exponen cuáles son los marcos legales a los que estas herramientas tecnológico-educativas deberían rendir cuentas o qué leyes deben conocer los desarrolladores y usuarios finales en cuanto al uso de tecnología en el aula. No es intención ofrecer una visión exhaustiva, pero sí suficiente, para conocer qué marco legal atiende a la recolección de datos en vigor y que afecta de forma directa al tratamiento de datos personales de los estudiantes en procesos de *Learning Analytics*. Se termina con un estudio de la población educativa española en cuanto al conocimiento y aplicación de estas leyes en instituciones educativas.

La Constitución Española de 1978 prevé en su artículo 18.4 que el legislador limitará el uso de la informática para proteger los derechos fundamentales de los ciudadanos. Unos años después, y coincidiendo con los Juegos Olímpicos de Barcelona en el 1992, se aprueba la Ley Orgánica de Regulación del Tratamiento Automatizado de los Datos de carácter personal (LORTAD) (Sánchez Bravo, 1994), que define por primera vez los datos de carácter personal y la identificación del afectado en este territorio.

La Ley Orgánica de Protección de Datos o Ley 15/1999 (LOPD) (Mañas, 2005) aprobada el 13 de diciembre de 1999, deroga la LORTAD y afecta a todos los datos que hacen referencia a personas físicas registradas sobre cualquier soporte, ya sea informático o no. La LORTAD solo ha estado activa 7 años, aunque marca un hito en la protección de datos personales al ser la primera ley aprobada de este carácter y además menos ambigua.

Este cambio legislativo parece ser de menor calidad con apertura a ambigüedades, puesto que la LORTAD tiene mayor riqueza, concreción y excepciones. En cambio, la LOPD no contiene ni exposición de motivos, aunque, entre otras cosas, describe su contenido, indica su objeto y finalidad. Cabe destacar que la LOPD es fruto de la complicada adaptación de la LORTAD a la Directiva 95/46/CE de protección de datos,

que finalmente creó ley nueva, la misma LOPD, siendo sustituida por la LOPDGDD (Boletín Oficial del Estado, 2018). Se complementa este contexto legal con el Real Decreto-ley 14/2019 (Boletín Oficial del Estado, 2019), de 31 de octubre, por el que se adoptan medidas urgentes por razones de seguridad pública en materia de administración digital, contratación del sector público y telecomunicaciones.

La Directiva 95/46/CE constituye el texto de referencia a nivel europeo en cuestiones de protección de datos personales (Baturones, 1998). En resumen, define un marco que permite la libre circulación de datos personales dentro de la Unión Europea y solicita que en cada Estado miembro exista un organismo nacional independiente supervisor de actividades relacionadas con el tratamiento de datos personales. No obstante, la existencia de una autoridad independiente que vele por un derecho de tráfico lícito está prevista en el Convenio 108 del Consejo de Europa, de 1981, el primer texto internacional sobre la materia.

II.6.1. Reglamento General de Protección de Datos

El Reglamento General de Protección de Datos (RGPD) (EP and the CEU, 2016) deroga la Directiva 95/46/CE y la LOPDGDD (Boletín Oficial del Estado, 2018) sustituye a la LOPD. Es interesante conocer los principios del RGPD e incluso cómo afecta a la transferencia internacional de datos. Se presenta el contexto en las siguientes secciones.

II.6.1.1. Principios del RGPD

Junto al RGPD se imponen nuevos principios, nuevas obligaciones para empresas, administraciones y otras entidades y nuevos derechos para los ciudadanos. Se presentan los nuevos principios de:

- Responsabilidad.
- Protección de datos por defecto y desde el diseño.
- Transferencia.

También se presentan nuevas obligaciones para empresas, administraciones y otras entidades de manera que se deben:

- Establecer un Delegado de Protección de Datos (DPD).

- Realizar evaluaciones de impacto sobre la privacidad.
- Ofrecer garantías adicionales para las transferencias internacionales de datos.
- Anular la obligación de inscribir los ficheros de datos y se incrementan las sanciones.

A los ciudadanos se les confiere nuevos derechos como:

- El de transparencia e información.
- Derecho al consentimiento inequívoco.
- Derecho de supresión (al olvido).
- Derecho a la limitación del tratamiento.
- Derecho a la portabilidad de los datos.
- Derecho a denuncias.
- Derecho a indemnizaciones por tratamiento ilícito de datos privados.
- Derecho al canon a la contestación de los ejercicios del derecho de acceso.

Este nuevo reglamento RGPD está vigente en todos los países de la Comunidad Europea. Por dicho motivo, podríamos asumir una cierta tranquilidad, aunque, no obstante, cómo se tratan los datos recolectados a nivel mundial es algo en lo que debe prestarse especial atención. Un claro ejemplo es la relación Unión Europea con Estados Unidos. Las leyes que regulan la transferencia de datos entre continentes levantan recelos en cuanto quién tiene acceso una vez los datos han aterrizado en los países destinatarios. De la misma forma es interesante saber si los datos deben residir en el país de su generación o por el contrario puede circular libremente bajo unas directrices concretas. Conocer cuál es el procedimiento, leyes y posibles fisgones es de antemano un conocimiento útil en el momento de elegir herramientas tecnológicas que afecten a datos educativos.

II.6.1.2. Leyes actuales aplicables

La LOPDGDD (Boletín Oficial del Estado, 2018; Pauner-Chulvi & Viguri Cordero, 2018) es por lo tanto la ley que ahora está vigente y que rige cualquier «transacción datarial» en el contexto educativo. Cualquier herramienta tecnológica educativa española que recolecte datos personales deberá cumplir los puntos definidos en esta ley. Como agentes educadores -directivos, profesores, padres, etc.- es imperativo preocuparse por

mantener un orden legal ante el uso de herramientas tecnológicas educativas. Por ahora la denuncia es el único instrumento para detener posibles tratamientos de datos ilícitos -en materia tecnológica no ha aparecido ninguna herramienta que asegure una seguridad y privacidad de datos educativos-. Pedir el cumplimiento de la LOPDGDD es el primer paso que se debe dar como agente educativo.

II.6.2. Protección y privacidad de datos Unión Europea- Estados Unidos

La transferencia de datos entre la Unión Europea y Estados Unidos requiere de un análisis que nos sitúe en contexto, puesto que ha estado regulada por distintas leyes a lo largo del tiempo.

II.6.2.1. Safe harbor

Un *safe harbor* (Weiss & Archick, 2016) es un reglamento que especifica que cierta conducta no viola una norma. Este tipo de reglamento o disposición de ley acostumbra a encontrarse en relaciones entre países debido a una ambigüedad lingüística. De esta forma se procura reducir la incertidumbre legal existente.

La Directiva europea 95/46/CE es un ejemplo de una ley *safe harbor*. Prohíbe a estados europeos compartir información de carácter personal a países que estén considerados de un nivel de protección por debajo de unos estándares de calidad. La posibilidad de establecer excepciones está abierta si los países deciden cumplir, y lo hacen, con los principios establecidos en el *safe harbor* de esta directiva.

II.6.2.2. Principios internacionales safe harbor

La Unión Europea define en la Directiva 95/46/CE en cooperación con el Departamento de Comercio de Estados Unidos unos principios que las organizaciones de este último país deben cumplir en materia de protección de datos personales, con el objeto de ser consideradas seguras y poder hacer transferencia de datos. De esta forma se evitan pérdidas o filtración no autorizada de datos personales. La emisión de un certificado anual indica la capacidad de las organizaciones de Estados Unidos de operar

en calidad de ente cumplidor de siete principios suficientes para garantizar una protección de datos personales:

- Información: debe informarse de que los datos personales recogidos serán tratados para la funcionalidad por la cual se recogen.
- Elección: existe el derecho de cancelación, a la oposición a datos recogidos una vez recabados y a la oposición de cesión o transferencia a terceros.
- Transferencia progresiva: solo pueden transferirse datos a terceros que cumplan con un adecuado nivel de cumplimiento de protección de datos.
- Seguridad: deben cumplirse unos criterios de seguridad para evitar pérdida de datos y filtraciones no autorizados.
- Integridad de los datos: debe asegurarse que los datos recabados sean correctos y relevantes para el propósito por los cuales fueron recabados.
- Acceso: debe existir el derecho de acceso, rectificación o eliminación de los datos recabados.
- Ejecución: debe garantizarse el cumplimiento de los 7 principios destinando los medios y recursos necesarios.

II.6.2.3. Unión Europea-Estados Unidos Escudo de privacidad

El reglamento y principios del *safe harbor* se anula en octubre de 2015 por el Tribunal de Justicia de la Unión Europea en la sentencia referida al caso Schrems (Ojanen, 2016). Maximillian Schrems hace una reclamación basándose en que sus datos personales se comprometen ya que Facebook Ireland Ltd transfiere a Estados Unidos los datos personales de sus usuarios, incluidos los suyos, además de conservarlos en los servidores ahí situados.

Se destaca que el sello del escudo de privacidad se expide por las propias organizaciones de Estados Unidos avaladas por el Departamento de Comercio de este país. Gracias a la denuncia de Schrems la Comisión Europea anula el *safe harbor* del 2015 para dar paso a otras directivas con el objeto de dar protección y privacidad a los datos personales de los ciudadanos europeos.

En febrero de 2016 la Comisión Europea y el gobierno de Estados Unidos llegan a un nuevo acuerdo en el que se establecen unos nuevos principios bajo el nombre Escudo de Privacidad (Weiss & Archick, 2016).

II.6.2.4. Principios del Escudo de Privacidad

El Escudo de Privacidad establece una serie de derechos para los ciudadanos y obligaciones a empresas para asegurar una protección de datos personales (AGPD, 2016). Dichos derechos son:

- Derecho a ser informado.
- Limitación en el uso de sus datos para diversos fines.
- Minimización de los datos y obligación de guardar los datos únicamente durante el tiempo necesario.
- Obligación de asegurar los datos.
- Obligación de proteger los datos si se transfieren a otra empresa.
- Derecho de acceso y rectificación de sus datos.
- Derecho a presentar una reclamación y a obtener reparación, y reparación en caso de acceso por parte de autoridades estadounidenses.

Es interesante comprobar que hay una última directriz directamente relacionada con las autoridades estadounidenses. La intención de este principio es limitar el acceso de estas autoridades a los datos personales de ciudadanos europeos recolectados por organizaciones privadas. Solo puede existir un acceso posible si hay un interés público, relacionado con la seguridad del país o la aplicación de sus leyes.

Tomando los principios del escudo de privacidad, las autoridades públicas de dicho país deben informar a los afectados en caso de acceder o al menos ofrecer la posibilidad de saber que se está accediendo y por la acción en concreto. En definitiva, es un principio muy genérico en el que no se termina de blindar el acceso de estados unidos a los datos personales de ciudadanos europeos.

II.6.3. Conocimiento de las leyes educativas

Las leyes educativas pueden ser complejas a ojos de ciudadanos no expertos en términos legales, e incluso para instituciones educativas que están más acostumbradas a ellas. Por dicho motivo las distintas agencias de protección de datos españolas ponen a disposición distintos documentos facilitadores de su implantación (Ramírez Gómez, 2018). No obstante, en dichos documentos no se especifica el nivel de adopción en instituciones educativas y/o roles educativos tras las leyes promulgadas entre 2018 y 2019, datos fundamentales para comprender un aspecto más del problema y ahondar en el desarrollo del estado del arte. Al mismo tiempo, la comprensión y percepción por parte de educadores de los peligros relacionados tanto con Internet-Web como con el uso de herramientas digitales educativas permite ahondar en aspectos relevantes de la investigación a modo de descripción del problema. En este sentido, durante septiembre del 2018 se realiza una primera encuesta a 117 participantes y una segunda en septiembre del 2019 a 154 participantes. Ambos resultados se comparan a modo de conclusión.

La construcción de dos cuestionarios adaptados a los cambios legislativos, uno para 2018 y otro para 2019 (ver Tabla 5 y Tabla 6), sigue los siguientes objetivos:

- Dilucidar el nivel de conocimiento de las leyes de protección de datos por parte de roles educativos de España.
- Entender cómo el transcurso del tiempo moldea este conocimiento.
- Conocer qué percepción tienen los roles educativos en relación con el uso de herramientas educativas que tratan datos.
- Conocer posible transferencia de conocimiento sobre leyes desde los centros a los profesores.

Tabla 5 Instrumento de recogida de nivel de conocimiento de leyes educativas por parte de roles educativos de España 2018

Pregunta	Tipo
P20181. Conozco la LOPD/RGPD	Likert 5 puntos
P20182. Aplico la normativa LOPD/RGPD en el aula	Likert 5 puntos

P20183. La privacidad de los alumnos es muy importante	Likert 5 puntos
P20184. Salvaguardar la identidad digital de los alumnos es muy importante	Likert 5 puntos
P20185. Me preocupan los datos que recogen las herramientas educativas	Likert 5 puntos
P20186. Leo las políticas de privacidad de las herramientas educativas para conocer cómo tratan los datos educativos	Likert 5 puntos
P20187. Soy consciente de los peligros que conlleva utilizar herramientas que no cumplan la LOPD/RGPD	Likert 5 puntos
P20188. Salvaguardar los datos de los alumnos ante usos indebidos es muy importante	Likert 5 puntos
P20189. Mi centro me ha indicado directrices sobre política y tratamiento de datos educativos	Likert 5 puntos

Tabla 6 Instrumento de recogida de nivel de conocimiento de leyes educativas por parte de roles educativos de España 2019

Pregunta	Tipo
P20191. Conozco la LOPDGDD o RGPD	Múltiples
P20192. Conozco qué son las leyes de protección de datos personales	Likert 5 puntos
P20193. Aplico las normativas de protección de datos personales en el aula	Likert 5 puntos
P20194. La privacidad digital de los alumnos es muy importante	Likert 5 puntos
P20195. Salvaguardar la identidad digital de los alumnos es muy importante	Likert 5 puntos
P20196. Salvaguardar los datos de los alumnos ante usos indebidos es muy importante	Likert 5 puntos
P20197. Me preocupa qué datos recogen las herramientas educativas digitales	Likert 5 puntos
P20198. Me preocupa cómo usan los datos recogidos por las herramientas educativas digitales	Likert 5 puntos
P20199. Leo las políticas de privacidad de las herramientas educativas digitales para conocer cómo tratan los datos educativos recolectados	Likert 5 puntos

P201910. Soy consciente de los peligros que conlleva utilizar herramientas educativas digitales que no cumplan las leyes de protección de datos personales	Likert 5 puntos
P201911. Mi centro me ha indicado directrices sobre política y tratamiento de datos educativos relativas a las leyes de protección de datos personales	Likert 5 puntos

II.6.3.1. Resultados

En los dos cuestionarios se realizan las mismas preguntas, a pesar de que en el cuestionario del 2019 hay dos más. Esto es debido a que el cuestionario del 2019 va relacionado con una línea paralela para conocer aspectos distintos sobre las leyes de protección de datos personales. En la Tabla 7 se hace el mapeo de las preguntas entre el cuestionario del 2018 y el 2019.

Tabla 7 Mapeo de preguntas entre cuestionarios del 2018 y 2019

Pregunta 2018	Pregunta 2019
P20181. Conozco la LOPD/RGPD	P20191. Conozco la LOPDGDD o RGPD
P20182. Aplico la normativa LOPD/RGPD en el aula	P20193. Aplico las normativas de protección de datos personales en el aula
P20183. La privacidad de los alumnos es muy importante	P20194. La privacidad digital de los alumnos es muy importante
P20184. Salvaguardar la identidad digital de los alumnos es muy importante	P20195. Salvaguardar la identidad digital de los alumnos es muy importante
P20188. Salvaguardar los datos de los alumnos ante usos indebidos es muy importante	P20196. Salvaguardar los datos de los alumnos ante usos indebidos es muy importante
P20185. Me preocupan los datos que recogen las herramientas educativas	P20197. Me preocupa qué datos recogen las herramientas educativas digitales
P20186. Leo las políticas de privacidad de las herramientas educativas para conocer cómo tratan los datos educativos	P20199. Leo las políticas de privacidad de las herramientas educativas digitales para conocer cómo tratan los datos educativos recolectados

P20187. Soy consciente de los peligros que conlleva utilizar herramientas que no cumplan la LOPD/RGPD	P201910. Soy consciente de los peligros que conlleva utilizar herramientas educativas digitales que no cumplan las leyes de protección de datos personales
P20189. Mi centro me ha indicado directrices sobre política y tratamiento de datos educativos	P201911. Mi centro me ha indicado directrices sobre política y tratamiento de datos educativos relativas a las leyes de protección de datos personales

En la Tabla 8 se muestran los resultados negativos de cada una de las preguntas comparadas entre cuestionarios. Se quiere conocer el desconocimiento en relación con las leyes sobre protección y confidencialidad de datos que aplican en contexto educativo.

Tabla 8 Resultados comparativos de las preguntas del cuestionario 2018 y 2019

Preguntas	Resultados 2018	Resultados 2019
P20181 y P20191	Un 17,9% no están de acuerdo o están algo en desacuerdo en conocer las leyes de protección de datos personales vigentes	Un 31% no conocen las leyes de protección de datos personales vigentes
P20182 y P20193	Un 17% no están de acuerdo o están algo en desacuerdo en aplicar las leyes de protección de datos personales en el aula	Un 7,7% no están de acuerdo o están algo en desacuerdo en aplicar las leyes de protección de datos personales en el aula
P20183 y P20194	Un 0,85% están algo en desacuerdo en considerar que la privacidad de los estudiantes es muy importante	Un 1,2% no están de acuerdo o están algo en desacuerdo en considerar que la privacidad de los estudiantes es muy importante
P20184 y P20195	Un 0% están algo en desacuerdo en considerar que salvaguardar la identidad digital de los estudiantes es muy importante	Un 2,5% no están de acuerdo o están algo en desacuerdo en considerar que salvaguardar la identidad digital de los estudiantes es muy importante

P20188 y P20196	Un 2,6% están algo en desacuerdo en considerar que salvaguardar los datos de los estudiantes ante usos indebidos es muy importante	Un 3,2% no están de acuerdo o están algo en desacuerdo en considerar que salvaguardar los datos de los estudiantes ante usos indebidos es muy importante
P20186 y P20199	Un 39,3% no están de acuerdo o están algo en desacuerdo en que lean las políticas de privacidad de las herramientas educativas para conocer cómo tratan los datos educativos	Un 33,1% no están de acuerdo o están algo en desacuerdo en que lean las políticas de privacidad de las herramientas educativas para conocer cómo tratan los datos educativos
P20187 y P201910	Un 14,5% no están de acuerdo o están algo en desacuerdo en que sean conscientes de los peligros que conlleva utilizar herramientas que no cumplan las leyes de protección de datos	Un 5,8% no están de acuerdo o están algo en desacuerdo en que sean conscientes de los peligros que conlleva utilizar herramientas que no cumplan las leyes de protección de datos
P20189 y P201911	Un 37,1% no están de acuerdo o están algo en desacuerdo en que su centro les haya indicado directrices sobre política y tratamiento de datos educativos	Un 32,5% no están de acuerdo o están algo en desacuerdo en que su centro les haya indicado directrices sobre política y tratamiento de datos educativos

La Figura 12 muestra un resumen del análisis de la Tabla 8:

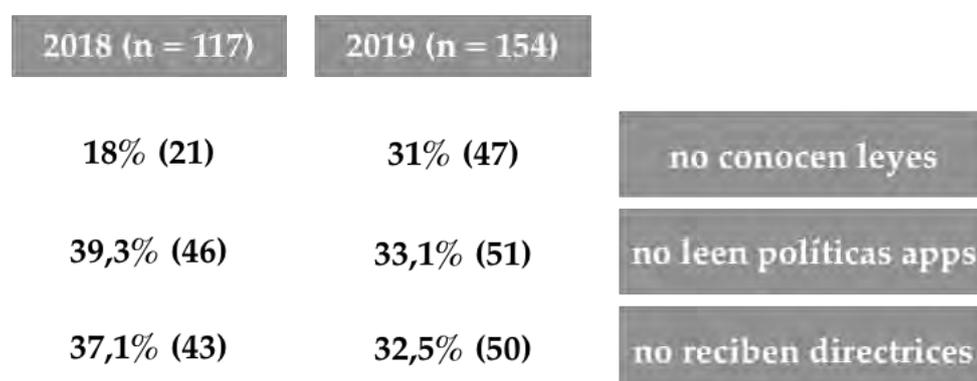


Figura 12 Resultado encuesta longitudinal 2018-2019 con relación a el conocimiento leyes de protección de datos en el aula. Elaboración: propia.

II.6.3.2. Análisis

Los datos registrados en las encuestas dibujan una realidad en la que aún parte de roles educativos en España desconocen total o parcialmente las leyes de protección de datos personales. En este sentido en el 2018 solamente un 17,9% de los encuestados no están de acuerdo o están algo en desacuerdo en conocer las leyes, en contraposición del 31% del 2019.

Los equipos directivos transfieren menos aspectos legales a sus docentes. Un 37,1% en el 2018 no están de acuerdo o están algo en desacuerdo en que su centro les haya indicado directrices en contraposición del 32,5% en el 2019. A pesar de este descenso, la aplicación de las leyes impacta en las aulas, puesto que las leyes pasan de no aplicarse total o parcialmente por un 17% de los encuestados en el 2018 al 7,7% en el 2019. Se deduce que los roles educativos desconocen detalles de las leyes, pero siguen procedimientos dictaminados por los centros.

Es importante notar que la percepción de la privacidad y seguridad de los alumnos es levemente menos importante en el 2019 que en el 2018. Este hecho no concuerda con que el 5,8% de los encuestados en el 2019 no conozca los peligros de usar herramientas digitales en comparación con el 14,5% del 2018. A pesar de esta despreocupación, el porcentaje de roles educativos que no están de acuerdo o están algo en desacuerdo en leer las políticas de privacidad de las herramientas educativas es levemente menor en el 2019 (33,1%) que en el 2018 (39,3%).

II.7. Blockchain

Los miedos y recelos generados alrededor del uso de *Learning Analytics* ponen de manifiesto una clara desconfianza ante esta aproximación analítico-educativa (ver II.4.4 Miedos y recelos: una cuestión delicada). Para eliminarla deben proponerse soluciones tecnológicas efectivas ante la gestión de la identidad de los estudiantes y la privacidad de los datos recolectados. La privacidad de datos y el anonimato de la identidad de los alumnos es el pilar más importante en el que trabaja la presente investigación para

establecer una confianza mínima ante la integración y normalización de procesos de *Learning Analytics* en instituciones educativas.

La necesidad de proteger identidades, metadatos y datos personales en distintos tipos de transacciones es imperativa para evitar fricciones, tanto desde la contratación de un profesional para realizar un trabajo en concreto pasando por la compraventa de una casa como hasta la expedición de un título universitario. Estas transacciones privadas y de fricción se han visto revolucionadas o posiblemente solucionables desde la irrupción de las llamadas criptomonedas, en concreto, del *Bitcoin*. Es la tecnología subyacente al *Bitcoin* la que ha hecho posible visionar cambios disruptivos en distintos contextos más allá de las criptomonedas.

En 2008 Satoshi Nakamoto (2008) propone la criptomoneda *Bitcoin* construida encima de su propuesta tecnológica *blockchain*. Con el uso de la tecnología *blockchain* se permite implementar un monedero distribuido gestionado por consenso, concepto que posibilita reinventar distintos contextos de negocio, por ejemplo, el del almacenamiento virtual. Otro ejemplo son las tecnologías derivadas tales como bloques programables con los que se pueden crear los llamados *Smart Contracts*. Estos contratos de ejecución automática permiten intercambiar dinero, propiedades, acciones o cualquier cosa de valor sin la necesidad de un intermediario. A la vez, y con tecnologías de encriptación, se puede preservar la privacidad y los metadatos de las transacciones. La implementación original de *blockchain* por parte de Nakamoto (2008) propone una solución a problemas de confianza entre pares capaz de eliminar intermediarios centralizados. Se abre un contexto de capas tecnológicas y criptográficas que interesados en su aplicación lanzan nuevas promesas de negocio.

Este contexto de tecnologías apiladas proyecta un futuro de soluciones tanto en los problemas de confianza en el uso de *Learning Analytics* como en cualquier otra situación educativa que requiera validación por consenso en un entorno de desconfianza. A continuación, se examina esta tecnología y se realiza una revisión sistemática de la literatura para comprender su impacto en el problema definido en la presente investigación.

II.7.1. Criptomonedas

Desde 1991 existe una preocupación en cómo certificar cuándo han sido creados o modificados documentos digitales del tipo imagen, video, texto o sonido. Las soluciones pasan por fusionar campos como la criptografía y servidores de marcas de tiempo (*timestamps*), pero sobretodo de distribuir la confianza, o lo que es lo mismo, utilizar un protocolo de consenso en la aceptación de transacciones.

La distribución de la confianza, o redes de usuarios distribuidos, es justamente la base de las criptomonedas. No obstante, en los años 90 no hay suficiente masa crítica de usuarios para llegar a una mínima capacidad de cálculo distribuida para que sean realmente efectivas. En el 2008 este contexto cambia hacia una realidad muy prometedora, aunque con algunas posibles limitaciones de futuro a subsanar.

Las criptomonedas existen desde 1998, cuando Wei Dai publica una descripción de un sistema de efectivo electrónico distribuido al que llama *b-money* (W. Dai, 1998). Al mismo tiempo Nick Szabo propone un esquema para una moneda digital descentralizada llamada "*bit gold*" que elimina la entidad central de validación de las transacciones (Szabo, 2008). Sin embargo, la idea de descentralizar el sistema y eliminar la entidad central de validación provoca problemas de fraude.

El principal problema con la moneda digital es la posibilidad de gastar la moneda dos veces. Esta situación de riesgo se conoce como *double-spend problem* (en castellano, el problema de doble gasto). Para evitar este problema, se necesita una agencia central para verificar que solo se realice un gasto de la moneda o *token* digital.

En las monedas digitales, y sin una agencia central de verificación, el fraude es más probable. Debe resolverse el problema de otra forma más socializada. Pasados diez años desde la aparición de la primera criptomoneda nadie es capaz de diseñar un nuevo esquema que resuelva el problema de forma efectiva. Satoshi Nakamoto lanza en 2008 una propuesta para *Bitcoin*, una nueva moneda digital basada en una cadena de hashes de pruebas de trabajo (*hash-proof-of-work chain*) (Nakamoto, 2008).

En el esquema propuesto por Nakamoto se define un libro contable distribuido por todos los usuarios de una red descentralizada. Por consiguiente, todos los usuarios

tienen una copia del libro contable. Para evitar el fraude, esta red valida por consenso todos los bloques de transacciones a añadir en el libro. La validación se realiza mediante pruebas de trabajo y criptografía. A este proceso de validación se le llama minar, y a los usuarios que participan se les recompensa regalándoles *tokens* de la criptomoneda. Así que, y en parte, la moneda *Bitcoin* se sustenta gracias al afán de los usuarios de obtener *tokens* de *Bitcoin* como recompensa por minar. Es por este motivo que no se comprende un sistema basado en tecnología *blockchain* sin una criptomoneda asociada. No se cree en la filantropía de las personas para ejecutar un proceso de minado sin recompensa.

II.7.2. Tecnología emergente

Blockchain, la tecnología detrás de la criptomoneda *Bitcoin*, tiene unas connotaciones exponenciales para los visionaros tecnológicos que iguala su potencial a la aparición de la imprenta, el motor o Internet, tres innovaciones que han cambiado el mundo de forma radical. Aún está por ver de qué manera esta tecnología de bloques concatenados cambiará nuestro mundo. Gartner muestra en su ciclo los distintos estadios en el que se encuentra la adopción de la tecnología *blockchain* en relación con los ámbitos en los que puede impactar (ver Figura 13). En algunos, los que están en el pico de las sobre expectativas, aún se está por ver qué impacto real va a tener y en otros cómo se empieza a estabilizar el uso de las criptomonedas y soluciones finales reales.

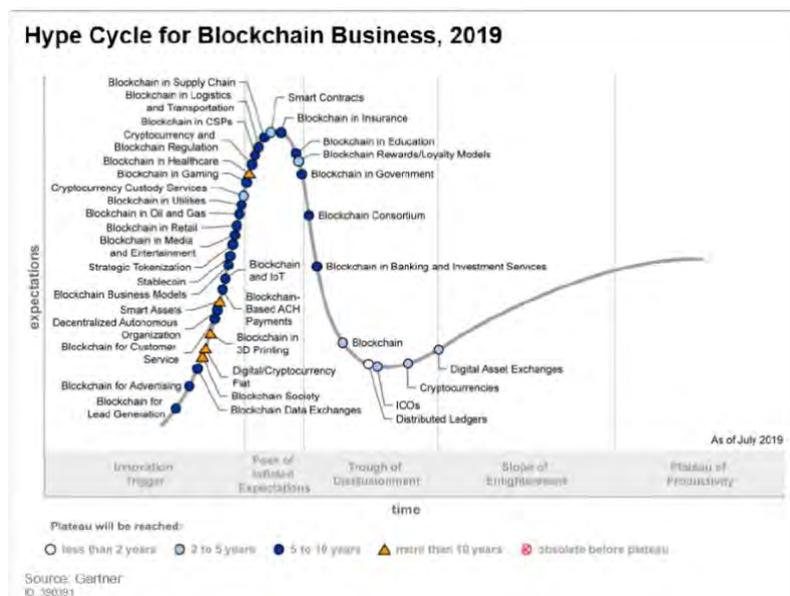


Figura 13 Ciclo de exageración de Gartner para negocios que usan tecnologías Blockchain. Fuente: (Gartner, 2019)

Las posibilidades de usar tal tecnología parecen ser ilimitadas. Justo ahora estamos aprendiendo a utilizarla y ya aparecen propuestas muy innovadoras como alojamiento en la nube descentralizado

Aunque las monedas virtuales e incluso las criptomonedas hace ya un tiempo que merodean por nuestro sistema económico, no ha sido hasta la coyuntura tecnológica actual que ha cobrado un sentido de utilidad. Esto es gracias a una evolución rápida tecnológica y sobre todo a la disponibilidad de capacidad de cálculo.

La tecnología que acostumbra a instaurarse es aquella más fácil de desarrollar. Sea *blockchain* o no la tecnología disruptiva definitiva, lo que en realidad se está replanteando en esta nueva revolución tecnológica es la confianza entre las entidades tradicionales y sus usuarios. Para entender el contexto de la presente investigación se requiere comprender las posibilidades de las criptomonedas para dar respuesta a los objetivos (ver I.2.1 Objetivos). A continuación, se detalla la relación entre educación, *Learning Analytics*, tecnología *blockchain* y tecnologías DLT (del inglés, *Distributed Ledger Technology*), para dar paso a una revisión sistemática de la literatura como parte de estado de la cuestión y marco para el desarrollo de posibles prototipos funcionales (ver I.2.3 Metodología).

II.7.3. Educación, *Learning Analytics* y *blockchain*

La aproximación antifraude propuesta por Nakamoto y su posible aplicación en el contexto educativo genera promesas para solucionar problemas en los que se requiere, por ejemplo, eliminar intermediarios. Ya se han ejecutado algunas soluciones piloto, como la gestión de certificados o almacenamiento de expedientes académicos. Otras no son más que ideas o conceptos sobre papel, no obstante, las posibilidades de uso de esta tecnología abre nuevas promesas en educación, a pesar de fuertes críticas en el sector académico (Adell et al., 2018):

- Mejorar la gestión de certificados.
- Proteger datos de las interacciones.
- Facilitar el compartir expedientes entre escuelas, universidades y empresas.
- Almacenar *e-portfolios* verificados.

- Gestionar propiedad intelectual.
- Mantener un histórico del comportamiento de los estudiantes.
- Gestionar acreditaciones.
- Nueva moneda interuniversitaria (educativa).
- Acciones automáticas según condiciones académicas.
- Analítica del aprendizaje automatizada.
- Identificación de estudiantes en entornos virtuales de aprendizaje.

La tecnología *blockchain* se define como una base de datos que promete la inmutabilidad de sus registros. Los registros son en realidad un conjunto de transacciones llamados bloques y enlazados unos con otros mediante *hashes* dependientes del bloque anterior. Esto significa que cualquier modificación obliga a reconstruir toda la cadena de bloques, hecho que comporta un coste de cálculo computacional desorbitado.

Esta característica de inmutabilidad aporta un nuevo significado en el contexto educativo, puesto que en el diseño de la tecnología *blockchain* el concepto de inmutabilidad se refiere a validez. Todo aquello que está en los bloques ha sido validado por consenso, existe un acuerdo previo y por consiguiente es válido y verdadero si dar lugar a dudas, malentendidos o disputas.

La investigación estudia esta serie de características de la tecnología *blockchain* con el objetivo de comprender sus posibilidades, y si es posible desarrollar una solución, para asegurar y blindar los datos educativos recolectados por herramientas de *Learning Analytics*.

II.7.4. Blockchain y DLT

En una arquitectura informática distribuida, tanto el almacenamiento como la computación son compartidos entre los usuarios o nodos de la red mediante conectividad punto-a-punto (P2P – del inglés *Peer To Peer*) (Conte de Leon, Stalick, Jillepalli, Haney, & Sheldon, 2017). En la tecnología DLT o DLS (del inglés, *Distributed Ledger System*), y bajo las condiciones adecuadas, es posible ofrecer una mayor disponibilidad de servicio y resiliencia. Este modelo distribuido ofrece algunas ventajas

y puede crear muchas oportunidades. Sin embargo, también crea grandes desafíos e interrogantes de investigación (Conte de Leon et al., 2017). Es el caso de los interesados en usar la tecnología *blockchain*, cuyas promesas son extensas pero su aplicación quizás no siempre necesaria (Wust & Gervais, 2018).

En cuestiones de arquitectura, se destaca que toda solución implementada con tecnología *blockchain* es una solución considerada como DLT, pero no toda solución considerada como DLT usa la tecnología *blockchain*. Por consiguiente, no todas las criptomonedas tienen una arquitectura basada en la tecnología *blockchain* cuyos bloques de transacciones son validados por protocolos de consenso de pruebas de trabajo (ver Figura 14).

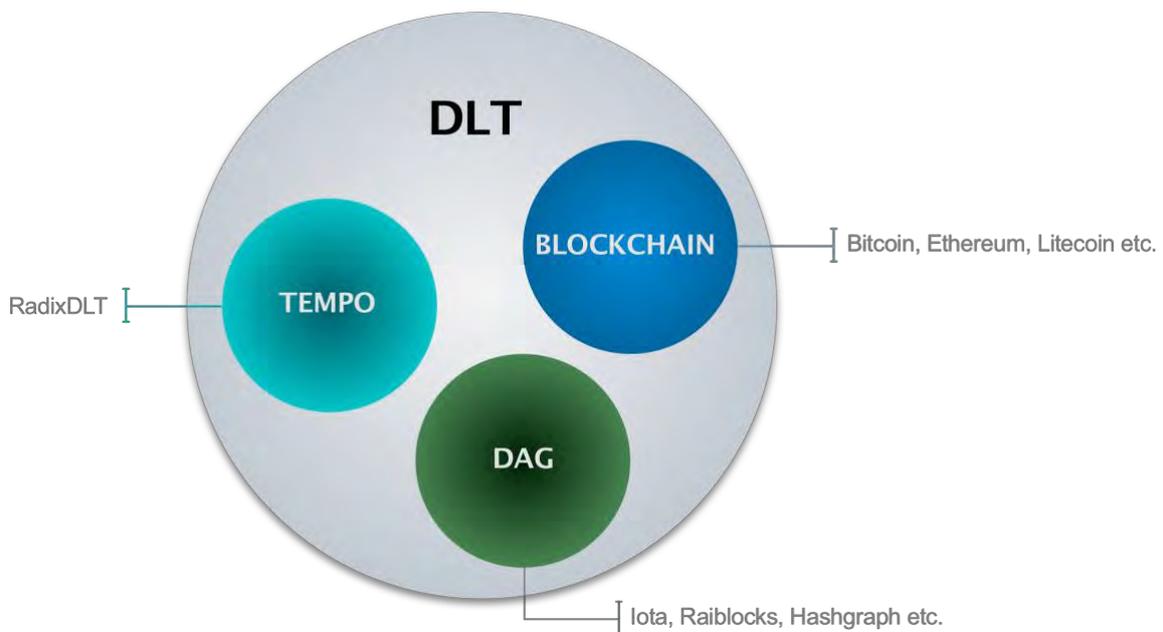


Figura 14 Organización de DLT, Blockchain y criptomonedas. Fuente: (Dexter, 2018)

Conte de Leon et al. (2017) ofrecen una descripción de la tecnología *blockchain* necesaria para diferenciarla de una DLT para evitar confusiones en la literatura. En este sentido, la tecnología *blockchain* es, en sus palabras, “...un método digital de registro de información capaz de registrar datos utilizando un diario con las siguientes características esenciales: ordenado, incremental, criptográficamente verificable hasta un bloque dado y digital”. En cambio, en palabras de Conte de Leon et al. (2017) “...un DLS, basado en DLT, es un sistema informático en el que un conjunto de procesos

informáticos que representan a agentes o usuarios conectados a una red digital funcionan de forma colaborativa sobre un conjunto de estructuras de datos de libro mayor distribuido”.

Las diferencias son importantes y en estas se basa la presente investigación para evitar confusiones en los términos usados. Asimismo, Conte de Leon et al. (2017) señalan como conclusión que se requiere una investigación multidisciplinar mucho más fundamental y práctica para garantizar la seguridad y la fiabilidad de los sistemas desarrollados antes de que se generalicen. Esta afirmación coincide con las predicciones de Gartner (Litan & Leow, 2019). Por consiguiente, cualquier solución desarrollada mediante tecnología *blockchain* debe considerar ambas conclusiones de los autores.

II.8. Revisión sistemática de la literatura

El uso y el análisis de los datos son ampliamente adoptados por todos gracias al uso de soluciones tecnológicas conectadas, dando lugar al movimiento *quantified self* (Lee, 2014). Además, las empresas basadas en datos tecnológicos utilizan los datos de los clientes para mejorar los servicios ofrecidos. De este modo, los datos se utilizan para ofrecer una mejor información a los clientes.

A nivel empresarial o incluso individual, los usuarios adquieren dispositivos y servicios conectados, generalmente por decisión propia (Bode & Kristensen, 2015). A través de estos dispositivos, las personas utilizan los servicios para enviar datos personales y biométricos a cambio de información accionable, como recomendaciones para la mejora de los deportes o advertencias relacionadas con el control de la salud (Meyer, Gurrin, Simske, Hermens, & Siek, 2014). Sin embargo, cuando se considera el contexto de la educación, podemos encontrar una diferencia principal: la voluntad de los usuarios.

Los datos de los clientes son la fuente de datos para los algoritmos de *Machine Learning*, técnicas de *Big Data* y análisis estadístico (Mayer-Schönberger & Cukier, 2013). Las agencias de servicios las utilizan para ofrecer información útil a los clientes, mejorar sus servicios y aprovechar el negocio. Los clientes son conscientes del uso de sus datos y aceptan a su voluntad los acuerdos de servicios sobre privacidad de datos y condiciones

de uso. Los usuarios manifiestan voluntariamente obtener resultados útiles a través del tratamiento de sus datos personales y sensibles. Esta voluntariedad no siempre es una opción disponible en el contexto de la educación, sobre todo cuando se integra *Learning Analytics* en el proceso de enseñanza.

Siemens introdujo el concepto de *Learning Analytics* en el 2010 (Siemens, 2010c) y ha evolucionado para abrazar contextos educativos tanto en línea como fuera de línea (Ochoa et al., 2017). Desde entonces, diferentes autores han identificado las posibilidades y los peligros de la aplicación de este método analítico (Chatti et al., 2012; Lupton & Williamson, 2017). Por un lado, un análisis del aprendizaje es capaz de mejorar y realzar el contexto de aprendizaje de una manera holística. *Learning Analytics*, tal y como lo define Erik Duval (2012), consiste en "recoger las huellas que los alumnos dejan atrás y utilizarlas para mejorar el aprendizaje". Esta definición, al igual que otras establecidas por diferentes autores (Next Generation, 2010; Siemens, 2010c), implica que las trazas de aprendizaje están asociadas a un amplio tipo de datos, incluido los personales (EP and the CEU, 2016). Estos rastros son necesarios para identificar y vincular al estudiante con los resultados del aprendizaje.

Learning Analytics procesa los datos de los estudiantes, incluso de los estudiantes menores de edad (Herold, 2014; Williamson, 2017b). El ciclo analítico consiste en recoger datos, almacenarlos durante largos períodos y utilizarlos para realizar análisis y visualizaciones (Amo & Santiago, 2017; Chatti et al., 2012). A mayor cantidad de datos, mejores resultados en el análisis. Este análisis puede ser descriptivo, predictivo e incluso prescriptivo, lo que implica la gestión, el tratamiento y la utilización de datos personales. Este contexto es muy sensible, a diferencia de los contextos individuales en los que el análisis se utiliza a voluntad. No está claro:

- ¿Cómo están utilizando los datos de los estudiantes las empresas de tecnología que dan servicio en educación y a quiénes realmente se les beneficia?
- ¿Cómo esto afectará a los estudiantes en un futuro a corto y largo plazo?
- ¿Qué nivel de privacidad o seguridad se aplica para proteger los datos de los estudiantes?

Por consiguiente, y en relación con lo expuesto, analizar datos educativos implica un contexto sensible y de fragilidad en la gestión y análisis de datos personales de los estudiantes, incluidos menores, en el que hay que maximizar las precauciones.

En esta revisión sistemática de la literatura (en inglés *Systematic Literature Review*, SLR) se explora la importancia de la protección y seguridad de los datos personales en el campo de la educación mediante las promesas emergentes de los interesados en usar la tecnología *blockchain*. Es importante entender las implicaciones de usar tecnologías emergentes, su relación con la sociedad y los riesgos legales derivados de sus distintos usos.

El presente apartado se organiza en distintas secciones en las que se expone:

- La metodología, sus aspectos, fases y pasos, para la revisión y el mapeo sistemático de la literatura.
- Los resultados extraídos del mapeo sistemático.
- Los resultados extraídos de la revisión sistemática de la literatura.
- El análisis y discusión sobre los descubrimientos encontrados en el proceso.
- Las amenazas a la validez del estudio.

II.8.1. Revisión y mapeo sistemáticos

Una revisión sistemática de la literatura (a menudo denominada revisión sistemática o revisión de la literatura) resume las evidencias existentes de un tema de investigación para presentar una evaluación de manera científica. La evidencia empírica es un tema importante para los investigadores de ingeniería de *software* para ayudar a identificar los vacíos en la investigación actual y proporcionar una base para identificar nuevas oportunidades de investigación. Kitchenham y Charters (2007) establecen una metodología para realizar revisiones rigurosas de la evidencia empírica actual a la comunidad de ingeniería de *software*. Por tanto, el propósito principal de una revisión sistemática para los investigadores de ingeniería de *software* es detectar, evaluar, comprender e interpretar los estudios disponibles en la literatura en relación con sus preguntas de investigación. El propósito de esta investigación tiene objetivos de

ingeniería de *software*, por tanto, se toma la aproximación de Kitchenham para la realización de la revisión sistemática.

Kitchenham y Charters (2007) introducen los estudios de mapeo sistemático como complemento de las revisiones sistemáticas. El objetivo de esta metodología es proporcionar una visión general de un área de investigación, identificar otras áreas adecuadas para llevar a cabo revisiones sistemáticas de la literatura y clasificar la literatura disponible para su uso en revisiones posteriores. Por consiguiente, el trabajo actual está organizado por las principales actividades propuestas por Kitchenham y Charters, tales como la planificación, realización y presentación de informes del estudio.

II.8.1.1. Revisión y planificación del mapeo

En el diseño de los procesos de revisión y mapeo, se identifican los diferentes objetivos, el protocolo de cumplimiento de objetivos y otros detalles relevantes anotados por Kitchenham y Charters. En las siguientes secciones, se establece una explicación de cada aspecto relevante.

El objetivo de utilizar el mapeo semántico es responder a las siguientes preguntas de mapeo (MQ, del inglés *Mapping Questions*) en relación con el problema detectado y el campo de estudio educación y *blockchain*:

- MQ1. ¿Cuántos estudios se han publicado a lo largo de los años?
- MQ2. ¿Quiénes son los autores más activos?
- MQ3. ¿Qué medios de publicación son los principales en la difusión de la investigación?
- MQ4. ¿En qué dominios se ha publicado?

El objetivo de utilizar la investigación sistemática es responder a las siguientes preguntas de investigación (RQ, del inglés *Research Questions*) en relación con el problema detectado y el campo de estudio educación y *blockchain*:

- RQ1. ¿Qué soluciones se han aportado en el campo de estudio?
- RQ2. ¿Qué problemas de seguridad presenta la tecnología *blockchain*?
- RQ3. ¿La tecnología *blockchain* cumple con el RGPD?

- RQ4. ¿Qué puede resolver la tecnología *blockchain* en relación con el problema?

Se define el alcance revisión bibliográfica en base al método PICOC (*Population, Intervention, Outcome and Context*) (Petticrew & Roberts, 2008). Sin embargo, este SLR no implica una fase de comparación.

- Población: Tecnología *blockchain* aplicada en la educación.
- Intervención: Soluciones que son desarrolladas, teorizadas o aplicadas para procesar datos educativos mediante tecnología *blockchain* de manera genérica, en procesos de *Learning Analytics* o de *Smart Contracts*.
- Comparación: No se planifica ninguna intervención de comparación.
- Resultados: Nivel de garantía de confidencialidad y seguridad de los datos e identidad de los estudiantes.
- Contexto: Entornos relacionados con la educación y la cadena de bloqueo, como las universidades.

II.8.1.2. Criterios inclusión y exclusión

Los criterios utilizados para incluir o excluir un trabajo se organizan en cuatro criterios de inclusión (IC, del inglés *Inclusion Criteria*) y cuatro criterios de exclusión (EC, del inglés *Exclusion Criteria*):

- IC1: La coincidencia presentada se aplica a los campos de la educación Y de *blockchain* (Y).
- IC2: La coincidencia presentada soporta procesos educativos O de *Learning Analytics* O de *Smart Contracts* O de seguridad O de privacidad O legales en relación con el RGPD (Y).
- IC3: Los trabajos se escriben en inglés (Y).
- IC4: Los artículos se publican en Revistas, Libros, Conferencias o Talleres revisados por pares.

Se establecen los siguientes criterios de exclusión:

- EC1: La coincidencia presentada no se aplica a los campos de educación O *blockchain* (O).
- EC2: La coincidencia presentada no soporta procesos educativos O de *Learning Analytics* O de *Smart Contracts* O de seguridad O de privacidad O legales en relación con el RGPD (O).
- EC3: Los trabajos no se escriben en inglés (O).
- EC4: Los artículos no se publican en Revistas, Libros, Conferencias o Talleres revisados por pares.

Se eligen los siguientes requisitos para las bases de datos electrónicas:

- La base de datos es capaz de utilizar expresiones lógicas o un mecanismo similar.
- La base de datos permite búsquedas de larga duración o búsquedas sólo en campos específicos de las obras.
- La base de datos está disponible para los investigadores (a través de las instituciones, a través de nuestra pertenencia a asociaciones como IEEE o ACM, que son responsables de algunas de las bases de datos utilizadas, etc.).
- La base de datos es una de las más relevantes en el área de investigación de interés de este proceso de mapeo: informática y educación.

La búsqueda se realiza en las siguientes bases de datos electrónicas:

- Digital ACM Library.
- Web of Science.
- IEEE Xplore.
- Springer Link.

II.8.1.3. Cadena de búsqueda

Para crear la cadena de búsqueda se realiza un proceso de identificación de los términos principales y significativos a partir de tres elementos esenciales: las preguntas de investigación, el PICOC, y las posibles variaciones ortográficas y sinónimas. En base a los términos identificados, se define una cadena de consulta utilizando los operadores

booleanos Y/O y el comodín (*) para encontrar cualquier palabra con sus diferentes terminaciones posibles (plural, singular, etc.). La cadena de consulta resultante es:

(blockchain Y education) O (blockchain Y "learning analytics) O (blockchain Y learning) O (blockchain Y "security issue*") O (blockchain Y "security challenge*") O (blockchain Y privacy) O (blockchain Y "privacy challenge*") O (blockchain Y "privacy challenge*") O (blockchain Y gdpr)

Esta cadena de consulta se adapta a cada una de las fuentes de búsqueda, resultando en una estructura de búsquedas muy exhaustiva por cada una de la base de datos electrónica (ver Tabla 9, Tabla 10, Tabla 11 y Tabla 12).

Tabla 9 Cadenas de búsqueda personalizadas para Digital ACM Library

Base de datos	Cadenas de búsquedas personalizadas	Resultados
Digital ACM Library	(+blockchain +education)	52
Digital ACM Library	(+blockchain +"learning analytics")	3
Digital ACM Library	(+blockchain +learning)	234
Digital ACM Library	(+blockchain AND +privacy)	276
Digital ACM Library	(+blockchain +"privacy challenge*")	2
Digital ACM Library	(+blockchain +"privacy issue*")	6
Digital ACM Library	(+blockchain AND +security)	408
Digital ACM Library	(+blockchain AND +"security challenge*")	3
Digital ACM Library	(+blockchain AND +"security issue*")	6
Digital ACM Library	(+blockchain AND GDPR)	2

Tabla 10 Cadenas de búsqueda personalizadas para Web of Science

Base de datos	Cadenas de búsquedas personalizadas	Resultados
Web of Science	TS=(blockchain AND education) Databases= WOS, CCC, DIIDW, KJ, MEDLINE, RSCI, SCIELO Timespan=All years Search language=English	95

Web of Science	TS=(blockchain AND learning) Research areas: (Education Educational Research)Databases= WOS, CCC, DIIDW, KJD, MEDLINE, RSCI, SCIELO Timespan=All years Search language=English	23
Web of Science	TS=(blockchain AND "learning analytics") Databases= WOS, CCC, DIIDW, KJD, MEDLINE, RSCI, SCIELO Timespan=All years Search language=English	4
Web of Science	TS=(blockchain AND "privacy challenge*") Databases= WOS, CCC, DIIDW, KJD, MEDLINE, RSCI, SCIELO Timespan=All years Search language=English	13
Web of Science	TS=(blockchain AND "privacy issue*") Databases= WOS, CCC, DIIDW, KJD, MEDLINE, RSCI, SCIELO Timespan=All years Search language=English	52
Web of Science	TS=(blockchain AND "security challenge*") Databases= WOS, CCC, DIIDW, KJD, MEDLINE, RSCI, SCIELO Timespan=All years Search language=English	32
Web of Science	TS=(blockchain AND "security issue*") Databases= WOS, CCC, DIIDW, KJD, MEDLINE, RSCI, SCIELO Timespan=All years Search language=English	60
Web of Science	TS=(blockchain AND GDPR) Databases= WOS, CCC, DIIDW, KJD, MEDLINE, RSCI, SCIELO Timespan=All years Search language=English	27

Tabla 11 Cadenas de búsqueda personalizada para IEEE Xplore

Base de datos	Cadenas de búsquedas personalizadas	Resultados
IEEE Xplore	(("Full Text & Metadata":blockchain) AND "All Metadata":education)	719
IEEE Xplore	(("Full Text & Metadata":blockchain) AND "All Metadata":"learning analytics")	19
IEEE Xplore	(("Full Text & Metadata":blockchain) AND "All Metadata":"privacy challenge*")	107

IEEE Xplore	("Full Text & Metadata":blockchain) AND "All Metadata":"privacy issue*")	430
IEEE Xplore	((("Full Text & Metadata":blockchain) AND "All Metadata":"security challenge*")	256
IEEE Xplore	("Full Text & Metadata":blockchain) AND "All Metadata":"security issue*")	638
IEEE Xplore	("Full Text & Metadata":blockchain) AND "All Metadata":"gdpr")	177

Tabla 12 Cadenas de búsqueda personalizada para Springer Links

Base de datos	Cadenas de búsquedas personalizadas	Resultados
Springer Links	blockchain AND education	996
Springer Links	blockchain AND learning Discipline=Education	28
Springer Links	blockchain AND "learning analytics"	30
Springer Links	blockchain AND "privacy challenge*"	49
Springer Links	blockchain AND "privacy issue*"	209
Springer Links	blockchain AND "security challenge*"	96
Springer Links	blockchain AND "security issue*"	333
Springer Links	blockchain AND gdpr	182

Con respecto a los resultados de las búsquedas, en general no se limitan por la fecha de publicación (la búsqueda se realiza incluyendo todos los artículos en el tiempo) ni se aplican otros filtros proporcionados por las bases de datos. En concreto, solo en Web of Science y Springer Link se acota la búsqueda relacionada con *blockchain* y privacidad para arrojar resultados en el campo educativo ante la gran cantidad de coincidencias presentadas.

II.8.1.4. Selección de literatura

Los trabajos recogidos se almacenan en una hoja de cálculo maestra. Posteriormente, se realiza un proceso de identificación, revisión, elegibilidad e inclusión (Moher, Liberati, Tetzlaff, & Altman, 2009) (ver Figura 15) reestructurado en las siguientes tres fases:

- Primera fase: Se eliminan duplicados y entradas erróneas para utilizarse en la segunda fase.
- Segunda fase: Se realiza una primera selección en base al título, al resumen y a los criterios de inclusión y exclusión definidos para la revisión sistemática de la literatura. Se realiza una evaluación rápida del contenido de aquellos trabajos no lo suficientemente valorables siguiendo el criterio inicial de esta primera fase. Los trabajos resultantes de este primer paso se almacenan en otra hoja de cálculo para empezar la tercera fase.
- Tercera fase: Se realiza una lectura de los artículos en profundidad y se analizan siguiendo las preguntas de la investigación. Los trabajos seleccionados se añaden a una última lista de verificación para una evaluación de calidad (ver Figura 15). El resultado se almacena en una hoja de cálculo definitiva. Se añade una referencia más resultado de la lectura de las referencias de los trabajos analizados.

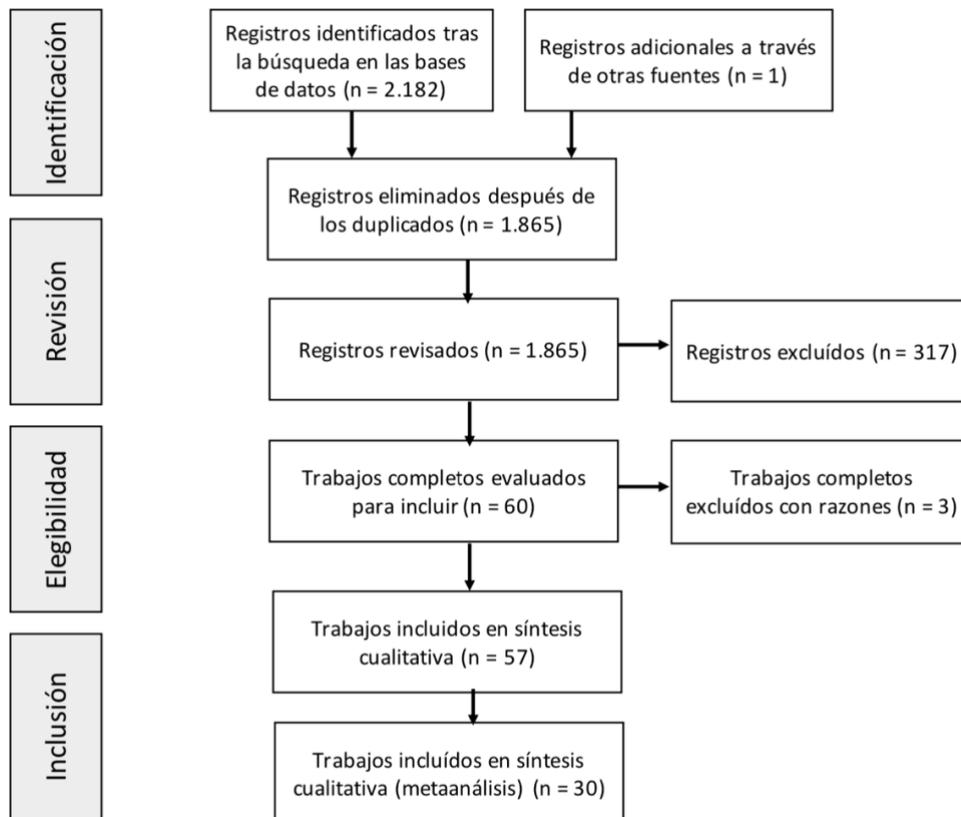


Figura 15 Pasos y resultados del proceso de revisión y mapeo. Informado como se propone en la declaración PRISMA. Fuente: (Moher et al., 2009)

La ejecución de las tres fases anteriores arroja los siguientes resultados:

1. Se ejecutan las cadenas personalizadas de búsqueda en las bases de datos. La ejecución arroja 2.182 trabajos que provienen de las bases de datos Digital ACM Library (473), Web of Science (178), IEEE Xplore (903) y Springer Links (628).
2. Se eliminan los trabajos duplicados. La eliminación deja 1.865 resultados por revisar (se incluyen aquellos de dudosa duplicidad).
3. Se revisan títulos y resúmenes. La revisión arroja 57 trabajos (3,21% de los trabajos únicos recuperados). Se incluye 1 referencia tras la revisión de referencias.
4. Tras la lectura del texto completo se seleccionan 30 trabajos (1,60% del total de trabajos considerados, 52,63% de los trabajos leídos).

Como se muestra en las directrices propuestas por Kitchenham y Charters (2007), se formula una lista de control de calidad para evaluar los estudios individuales y evitar la subjetividad. Estas listas de control son útiles para ayudar en el proceso de selección de los trabajos. La lista de verificación para la evaluación de la calidad elaborada se basa en la lista de verificación sugerida en (Kitchenham & Charters, 2007). Otros trabajos sobre revisiones sistemáticas y mapeo de la literatura (Cruz-Benito, García-Peñalvo, & Therón, 2019; Neiva, David, Braga, & Campos, 2016; Soomro et al., 2016) también personalizan sus listas de control de calidad basándose en las sugerencias dadas en (Kitchenham & Charters, 2007).

En la tercera fase la revisión, como se describe anteriormente, los trabajos se leen en su totalidad y su calidad se evalúa utilizando la lista de control de evaluación de calidad formulada (ver Tabla 13). La respuesta a cada una de las 10 preguntas se puntúa con 1 punto si la respuesta es "Sí", 0,5 puntos si la respuesta es "Parcial" o 0 si la respuesta es "No". Al utilizar este sistema, cada trabajo puede obtener una puntuación de 0 a 10 puntos. La marca del primer cuartil (Q1 = 7,5 puntos o más de 10 posibles) se utiliza como puntuación de corte para incluir un trabajo. Si un trabajo obtiene una puntuación inferior a 7,5, se excluye de la lista final para evitar trabajos de baja calidad de acuerdo con la lista de control de evaluación de calidad.

Tabla 13 Lista de control de evaluación de la calidad

Pregunta	Calificación
1. ¿Están claramente especificados los objetivos de investigación relacionados con la educación y <i>blockchain</i> ?	S / N / Parcial
2. ¿El estudio fue diseñado para lograr estos objetivos?	S / N / Parcial
3. ¿El enfoque <i>blockchain</i> está claramente descrito y justificado?	S / N / Parcial
4. ¿La investigación está respaldada por datos de algún tipo?	S / N / Parcial
5. ¿Se presentan soluciones sobre educación y <i>blockchain</i> ?	S / N / Parcial
6. ¿Se presentan soluciones acerca de dominios de privacidad, seguridad y regulación legal de datos personales?	S / N / Parcial
7. ¿Se ha explicado suficientemente la necesidad de la privacidad, seguridad o regulación legal de datos personales?	S / N / Parcial
8. ¿Los investigadores discuten algún problema de privacidad y seguridad de <i>blockchain</i> ?	S / N / Parcial
9. ¿Los vínculos entre datos, interpretación y conclusiones son claros?	S / N / Parcial
10. ¿Todas las preguntas de investigación se responden adecuadamente?	S / N / Parcial

Al aplicar la marca del primer cuartil solo se obtienen cuatro trabajos. Para cubrir un espectro más amplio se decide usar las puntuaciones que estén dentro del segundo cuartil (Q2 = 5 puntos o más de 10 posibles). Esta ampliación del espectro no entra en conflicto con la relevancia de la selección de trabajos, puesto que el campo de estudio abraza distintos dominios dentro del mismo que ayuda a alcanzar muchas perspectivas interesantes a incluir en la revisión de la literatura.

II.8.2. Resultados del mapeo sistemático

Se responde a las preguntas de mapeo (MQ) con los resultados del análisis de las publicaciones seleccionadas, en concreto 30 de 2.183. Todos los datos relativos a los trabajos publicados se almacenan y analizan en un documento Microsoft Excel, con las consecuentes hojas de cálculo internas para cada fase, paso analítico, revisión realizada,

tabla o figura creada. A continuación, se responde a las preguntas del mapeo sistemático en base a los resultados extraídos del análisis de los trabajos seleccionados.

En la Figura 16 se visualiza la respuesta a la primera pregunta sobre el mapeo MQ1. ¿Cuántos estudios se han publicado a lo largo de los años? En la revisión sistemática de la literatura no se ha restringido la búsqueda por años u otro criterio temporal. Los trabajos seleccionados durante el proceso de revisión y mapeo, desde el origen de la tecnología *blockchain* (Nakamoto, 2008), se publican entre el 2013 y el 2018, siendo el año 2018 el año con más publicaciones. Se identifica una tendencia al alza del interés en el campo tras el notable crecimiento de publicaciones a partir del 2016.

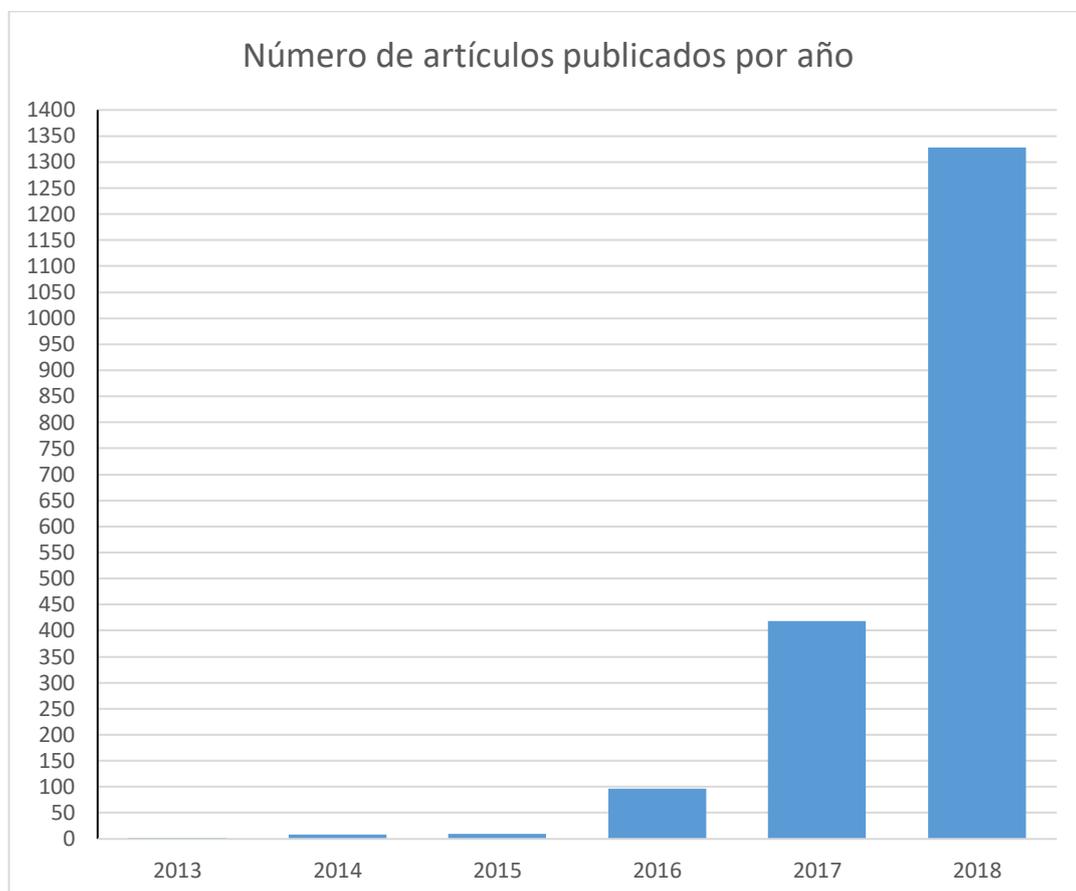


Figura 16 MQ1— Número de artículos publicados por año. Fuente: Elaboración propia

En relación con la segunda pregunta sobre el mapeo MQ2. ¿Quiénes son los autores más activos?, se identifica a todos los autores de los trabajos seleccionados. No se encuentra ningún autor con más de una publicación. Todos los 113 autores de los 30 trabajos

seleccionados en la revisión sistemática de la literatura pueden consultarse en la Tabla 14.

Tabla 14 Nombres de los autores y número de publicaciones para cada uno

Nombre	Total
Arthur Gervais; Ghassan O. Karame; Karl Wüst; Vasileios Glykantzis; Hubert Ritzdorf; Srdjan Capkun; Nelson Bore; Samuel Karumba; Juliet Mutahi; Shelby Solomon Darnell; Charity Wayua; Komminist Weldemariam; B. Duan; Y. Zhong; D. Liu; M. Apostolaki; A. Zohar; L. Vanbever; Yuqin Xu; Shangli Zhao; Lanju Kong; Yongqing Zheng; Shidong Zhang; Qingzhong Li; Daniel Drescher; Dai, Fangfang; Shi, Yue; Meng, Nan; Wei, Liang; Ye, Zhiguo; Bdiwi, Rawia; de Runz, Cyril; Faiz, Sami; Cherif, Arab Ali; Meng Han; Zhigang Li; Jing (Selena) He; Dalei Wu; Ying Xie; Asif Baba; Patrick Ocheja; Brendan Flanagan; Hiroaki Ogata; Alexander Mense; Markus Flatscher; N. Al-Zaben; M. M. Hassan Onik; J. Yang; N. Lee; C. Kim; J. C. Farah; A. Vozniuk; M. J. Rodríguez-Triana; D. Gillet; X. Gong; X. Liu; S. Jing; G. Xiong; J. Zhou; M. Turkanović; M. Hölbl; K. Košič; M. Heričko; A. Kamišalić; R. Arenas; P. Fernandez; A. Srivastava; P. Bhattacharya; A. Singh; A. Mathur; O. Prakash; R. Pradhan; S. Gilda; M. Mehrotra; G. Dima; A. Jitariu; C. Pisa; G. Bianchi; M. Conti; E. Sandeep Kumar; C. Lal; S. Ruj; J. Moubarak; E. Filiol; M. Chamoun; A. Soni; S. Maheshwari; Guang Chen; Bing Xu; Manli Lu; Nian-Shing Chen; Flanagan, Brendan; Ogata, Hiroaki; Millard, Christopher; Joshi, Archana Prashanth; Han, Meng; Wang, Yan; Turcu, Cristina; Turcu, Cornel; Chiuchisan, Iuliana; Sun, Han; Wang, Xiaoyue; Wang, Xinge; Pagallo, Ugo; Bassi, Eleonora; Crepaldi, Marco; Durante, Massimo; J. Bacon; J. Michels; C. Millard; K. Kuvshinov; I. Nikiforov; J. Mostovoy.	1

En relación con la tercera pregunta sobre el mapeo MQ3. ¿Qué medios de publicación son los principales en la difusión de la investigación?, se analizan los distintos tipos de publicaciones relacionados con los trabajos seleccionados.

Se observa en la Figura 17 una gran mayoría de trabajos seleccionados (21/30, 70%) son artículos publicados en conferencias. Por otro lado, los otros tipos de publicaciones encontrados son los artículos publicados en revistas (7/30, 23,33%) y capítulos de libro (1/30, 3,33%). Como complemento a este análisis de medios de difusión, se aporta el listado de los títulos de las publicaciones (ver Tabla 15). En el listado se muestra el número de las publicaciones seleccionadas y relativas al medio de difusión (ver Tabla 46 para más detalle relativo a los números y referencias), el título del medio de difusión y el índice H. El índice H se proporciona en base a cómo lo reporta la base de datos SCIMAGO (SJR). Se aporta este listado para exponer cuáles son los medios de publicación y cuáles son los más relevantes en la comunidad científica.

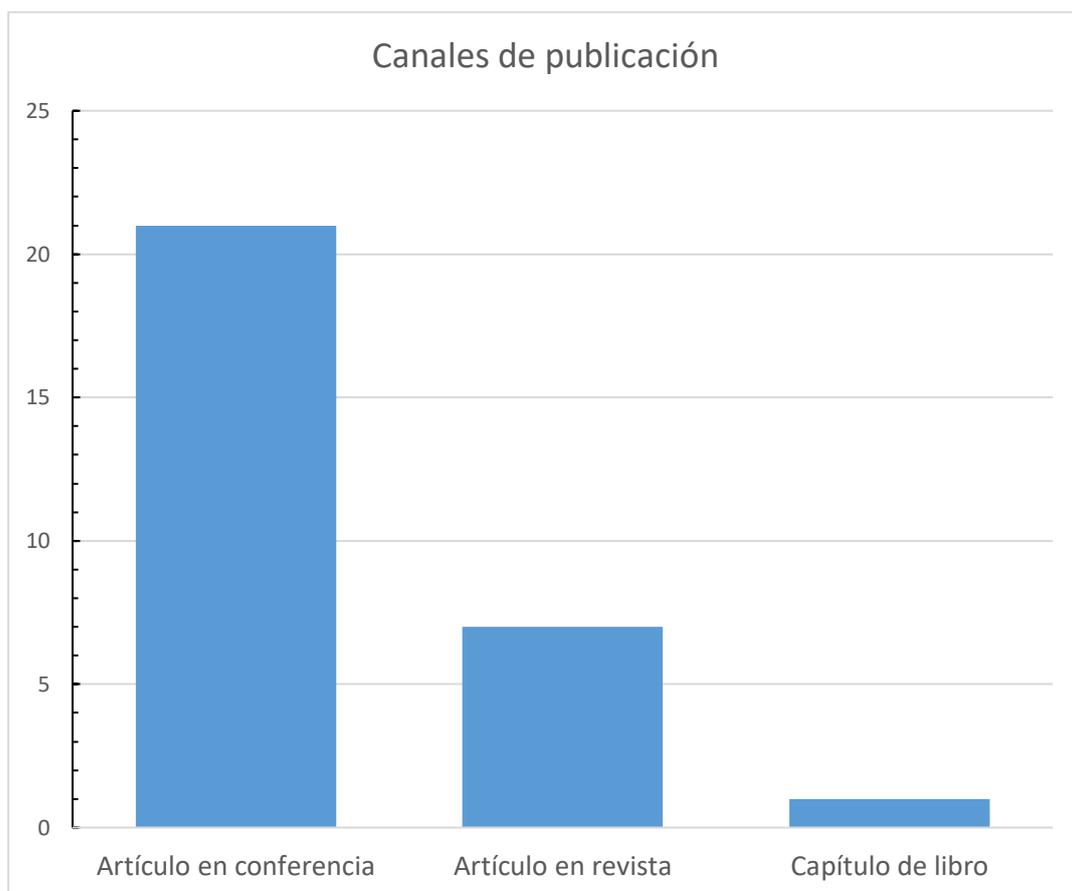


Figura 17 MQ3 Canal de publicación de los documentos seleccionados. Fuente: Elaboración propia

Tabla 15 Fuentes de publicación

Referencias	Nombre de la publicación	Índice H
[1]	ACM Conference on Computer and Communications Security	12
[9]	Annual SIG Conference on Information Technology Education	-
[6]	Blockchain Basics: A Non-Technical Introduction in 25 Steps	-
[14]	Chinese Automation Congress	-
[25]	Computer Law and Security Review	28
[5]	ICTAC Theoretical Aspects of Computing	-
[15]	IEEE Access	56
[20]	IEEE Communications Surveys & Tutorials	147
[8], [13]	IEEE International Conference on Advanced Learning Technologies	8
[16]	IEEE International Conference on Engineering, Technology and Innovation	3
[3]	IEEE International Conference on Parallel and Distributed Systems	35
[19]	IEEE International Forum on Research and Technology for Society and Industry	7
[22]	IEEE International Students' Conference on Electrical, Electronics and Computer Science	8
[21]	IEEE Middle East and North Africa Communications Conference	-
[4]	IEEE Symposium on Security and Privacy	97
[17]	International Conference on Advances in Computing, Control and Communication Technology	-
[18]	International Conference on Computer Communication and Informatics	5
[12]	International Conference on Computing, Electronics & Communications Engineering	-
[2]	International Conference on Information and Communication technologies for Development	11
[11]	International Conference on Information Integration and Web-based Applications & Services	10

[10]	International Conference on Learning Analytics and Knowledge	-
[7]	International Conference on Systems and Informatics	4
[27]	International Conference on Virtual Learning	-
[28]	International Journal of Emerging Technologies in Learning	15
[24]	Knowledge Management & E-Learning	18
[29]	Legal Knowledge and Information Systems	-
[26]	Mathematical Foundations of Computing	-
[30]	Richmond Journal of Law & Technology	-
[23]	Smart Learning Environments	-

En el caso de la pregunta MQ4: ¿En qué dominios se ha aplicado?, se obtienen los resultados a partir de las palabras clave de las publicaciones seleccionadas. Se observa que el dominio principal descrito en los trabajos tiene que ver con la tecnología *blockchain*, seguido por el dominio relativo a la educación. De las 145 palabras clave únicas utilizadas en los trabajos, la gran mayoría está relacionada con la palabra clave *blockchain*. En concreto, los dominios a los que se refieren las palabras clave de los trabajos (Figura 18), son: “Blockchain” (29/145, 20%), “Educación” (20/145, 13,79%), “Protección de datos” (16/145, 11,03% del total), “Criptomonedas” (14/145, 9,65% del total), “Distributed Ledger Technology” (5/145, 3,44%), “Seguridad” (5/145, 3,44%) y “Learning Analytics” (3/145, 2,06%).

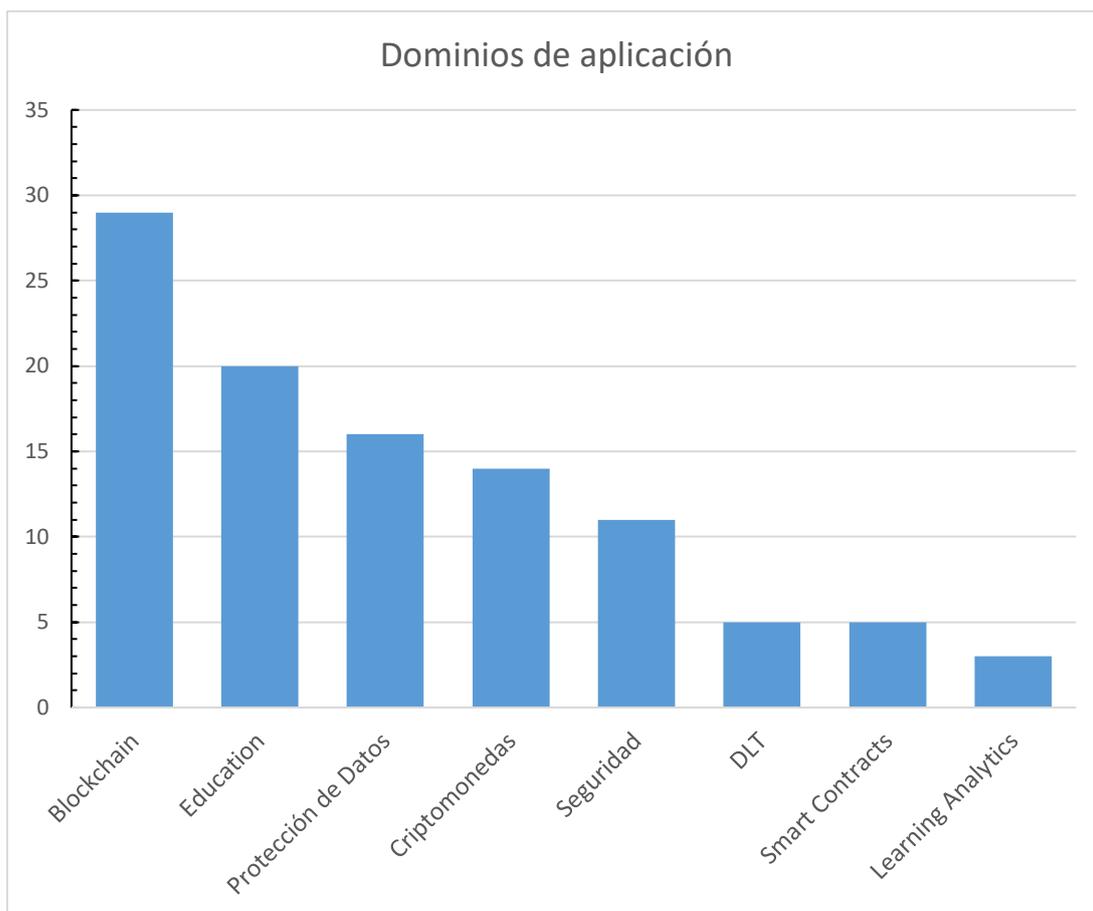


Figura 18 Dominios de aplicación. Fuente: Elaboración propia

La Tabla 16 resume los principales resultados obtenidos durante el proceso de mapeo. Aunque no se pueden considerar como una representación de todo el estado del arte entre la intersección de *blockchain* y educación, proporciona unos resultados concluyentes deliberados en la sección II.8.4 Análisis resumen de las soluciones propuestas.

Tabla 16 Resumen de los resultados del informe sobre el mapeo

Pregunta sobre el mapeo	Resultado
MQ1	Los artículos se publicaron entre los años 2016 y 2018
MQ2	Ninguno de los autores de los trabajos resultantes de la revisión sistemática de la literatura cuenta con más de una publicación
MQ3	La mayoría de los artículos han sido publicados en conferencias

MQ4	La gran mayoría de los trabajos tienen que ver con el dominio de <i>Blockchain</i> y Educación
-----	--

II.8.3. Resultados de la revisión sistemática

Se procede a dar respuesta a las preguntas de investigación (RQ) planteadas durante la planificación del trabajo de investigación. Para dar una respuesta ordenada se ha creado una taxonomía adicional que simplifica la comprensión del problema. La taxonomía que agrupa los distintos trabajos seleccionados se muestra en la Tabla 17 y parte de la lectura de cada uno de los trabajos. De cada respuesta se deduce el estado del campo de investigación (educación y *blockchain*) en consideración con las categorías de la Tabla 17.

Tabla 17 Taxonomía de trabajos seleccionados en la revisión sistemática de la literatura

Categoría	Trabajos
Educación	[2], [3], [5], [8], [9], [10], [13], [14], [15], [16], [17], [18], [19], [23], [24], [27], [28]
Seguridad	[1], [4], [6], [7], [11], [20], [21], [22], [26], [30]
Privacidad	[18], [20], [26], [30]
RGPD	[12], [25], [29], [30]

Se procede a dar respuesta secuencial a las preguntas de investigación para continuar con un resumen y análisis de la situación actual y unas reflexiones finales. La Tabla 18 a continuación relaciona los trabajos seleccionados con cada una de las preguntas de investigación a las que puede dar respuesta.

Tabla 18 Resumen de trabajos empleados para responder a las preguntas de investigación

Pregunta	Trabajos
RQ1	[2], [3], [5], [8], [9], [10], [13], [14], [15], [16], [17], [18], [19], [23], [24], [27], [28]
RQ2	[1], [4], [6], [7], [11], [20], [21], [22], [26], [30]
RQ3	[12], [20], [25], [29], [30]
RQ4	[7], [12], [18], [20], [24], [26], [30]

II.8.3.1. RQ1. ¿Qué soluciones se han aportado en el campo de estudio?

El campo de estudio de la presente investigación es considerado el binomio educación y *blockchain*. Conocer qué soluciones se aportan en educación basadas en tecnología *blockchain* es interesante para conocer qué problemas se pretenden resolver, qué problemas se creen que existen y requieren solución y, por supuesto, el estado de madurez de la cuestión.

El ámbito de la aplicación de la tecnología *blockchain* es muy variado y se encuentra actualmente en distintas fases de madurez (Iansiti & Lakhani, 2017) según el sector en el que se aplica (Litan & Leow, 2019). Por lo general, la tecnología *blockchain* se relaciona con “garantizar certificación”. Sin embargo, las promesas de esta tecnología van a veces más allá de una simple certificación temporal de los eventos (Wilkinson & Lowry, 2014). Así se muestra en las distintas publicaciones seleccionadas para dar respuesta a la primera pregunta de investigación “RQ1. ¿Qué soluciones se han aportado en el campo de estudio?”. Bore et al. (2017), investigadores en Kenia, proponen utilizar la tecnología *blockchain* para tomar mejores decisiones basadas en datos. Su solución se sitúa en el plano macro de educación, cuyos resultados se destinan a ser usados por gobiernos e instituciones no gubernamentales. Justifican la necesidad de usar un sistema único, global, interconectado y veraz debido a la imposibilidad de los distintos sistemas presentados hasta el momento, y en países en proceso de desarrollo, de recolectar información a tiempo real, objetiva y con resultados holísticos.

Para Bore et al. (2017), el hecho de que estos países en vías de desarrollo no tengan información estadística fiable, sin modificaciones, sin intenciones fraudulentas, actualizada a tiempo real y accesible de todo el sistema educativo, evita poder solucionar muchos de los retos que presenta educación. Ponen ejemplos de escuelas en Sudáfrica, donde los registros de los alumnos están centralizados en bases de datos relacionales que imposibilitan a ministros u otros departamentos gubernamentales recolectar la información de manera fidedigna, sin pérdida de datos o manipulaciones intermedias, y sin dedicar grandes esfuerzos y recursos. Del mismo modo, un alumno que quiere transferir sus registros de una institución educativa a otra debe hacerlo manualmente.

Como solución a estos problemas Bore et al. (2017) presentan la tecnología *blockchain* como un nuevo concepto en la gestión de datos y transacciones en negocios que permite establecer puntos de control fiables y una gestión adecuada para el cumplimiento de datos seguros (datos), la integridad de las transacciones (inmutabilidad) y asegurar que las transacciones sean identificables (no repudio). Ven una clara promesa en el uso de la tecnología *blockchain*, a través de algoritmos criptográficos, para asegurar que los datos de las distintas escuelas estén centralizados, se asegure su inmutabilidad y su acceso a tiempo real. Su propuesta reside en crear una solución basada en tecnología *blockchain* a la que llaman *School Information Hub* (Centro de Información Escolar). La arquitectura del sistema tiene entre otros módulos, un marco al que llaman *School Data Hub*, el *School Information Hub* y un almacenamiento de registros de escuela. El sistema se sustenta en una serie de *apps* que permiten recolectar datos demográficos, de documentos de personas y otros muy sensibles como biométricos. Junto a un preprocesador de datos, el cual incluye un módulo de *Learning Analytics*, realizan un conjunto de tratamientos analíticos para proveer *insights* acerca de estudiantes, profesores, escuelas o recursos para que los decisores tanto de niveles macro (gobiernos) como micro (escuelas y profesores).

Las transacciones ocurridas en un colegio, y diferenciadas por roles, son registros del tipo “inscripción en un curso” o “actualización de datos biométricos”. Para un profesor se puede registrar “alta de nuevo profesor” o “asignación de un curso”. Los registros con los datos de cada transacción se almacenan en los registros físicos o lógicos de la escuela. Probablemente, y no indicado en el trabajo, el almacenamiento de los registros de escuela se sustenta por una base de datos convencional. Al mismo tiempo, la automatización en este sistema juega un papel fundamental debido al gran conjunto de datos recolectado.

Duan, Zhong y Liu (2018) exponen en su trabajo cómo una propuesta de uso de la tecnología *blockchain* puede ayudar a superar el problema de la manca de efectividad en la verificación de las habilidades, conocimientos y logros de los estudiantes. El uso de automatizaciones, basadas en *Smart Contractas* dentro de su propuesta, argumentan que son necesarias para hacer el sistema funcional. La

originalidad de la propuesta reside en una doble cadena de bloques, una para gestionar las habilidades conseguidas por los estudiantes y otra para gestionar los cursos e incluso la calidad impartida por el profesor. Para el funcionamiento de su propuesta los autores definen tres aspectos fundamentales:

- El primer aspecto explica cómo se recolectan los datos de forma automatizada después de las comprobaciones pertinentes del profesor, puesto que solo se guardan en los bloques aquellos estudiantes cuyos resultados finales sean satisfactorios según umbral.
- El segundo aspecto expone qué información se almacenará en los bloques encadenados, dando cuenta que es información sensible del estudiante, y consiste en un registro con el valor total de los logros, el nombre del curso, el nombre del resultado, el peso del curso y otra información.
- El tercer aspecto muestra el protocolo de consenso de acreditación, basado en una fórmula de carácter cuantitativo y cualitativo.

Por consiguiente, con este sistema se pueden expedir micro-diplomas relacionadas con las competencias (habilidades, conocimientos y logros) de los estudiantes. Curiosamente se hicieron pruebas con el curso “seguridad de la información”. Cabe destacar que en el trabajo no se expone dónde se almacenan los registros de datos, aunque damos por entendido que se almacena en los *Smart Contracts* con un nivel adecuado de encriptación para asegurar la identidad y la confidencialidad de los datos de los alumnos.

Es probable que los autores del anterior trabajo no especifiquen niveles de encriptación debido a las dificultades que esto presenta para equilibrar la protección de datos de los estudiantes y la eficiencia en la búsqueda de información dentro de los bloques de transacciones. Esto es en lo que se centran los autores Xu et al. (2017). En su trabajo proponen un sistema que permita maximizar la eficiencia de búsquedas a la par que asegurar la privacidad de los datos de los alumnos en la gestión de certificaciones. Su solución presenta una plataforma de certificados en el que cualquier persona pueda certificar su autenticidad, identificarse y disponer de este cuando lo necesite. Para conseguirlo, proponen usar la tecnología *blockchain* para garantizar la seguridad de los

datos y la confianza en el sistema distribuido mediante la creación de bloques por cooperación. Es una aproximación interesante puesto que aboga por un protocolo de consenso no competitivo sino constructivo entre los usuarios de la red. Se destaca que, a diferencia de los trabajos encontrados, Xu et al. (2017) exponen con detalle cómo aseguran la privacidad de los datos disponibles en los bloques de datos definidos por la tecnología *blockchain* de manera que un acceso indebido evite el acceso a los datos extraídos.

Las implementaciones de la tecnología *blockchain* para garantizar la certificación de logros educativos son las más comunes. Turkanović et al. (2018) proponen el EduCTX, una solución más atomizada en el contexto de las certificaciones. El trabajo se enmarca en el Sistema Europeo de Transferencia y Acumulación de Créditos (ECTS, del inglés *European Credit Transfer and Accumulation System*), centrado en el estudiante y que marca la carga de trabajo para la consecución de los objetivos de un programa. Turkanović et al. (2018) proponen una solución que implementa la tecnología *blockchain* y de acceso privado para dar crédito certificado, además de ofrecer un sistema global de manera que también puedan aprovecharse compañías, instituciones y organizaciones. Con esta solución pretenden derribar barreras administrativas y de comunicación entre universidades, así como facilitar la certificación de los créditos obtenidos por los estudiantes.

Srivastava et al. (2019) ponen en entredicho el trabajo de Turkanović et al. (2018). Aseguran que falla el control de un estudiante cuando se registra en distintos cursos de un mismo profesor. Srivastava et al. (2019) presentan su solución a los problemas de certificación lenta y centralizada en papel o bases de datos digitales de instituciones educativas. Además, pretenden facilitar la transferencia de créditos entre universidades sin considerar las barreras legales y de políticas administrativas.

Arenas y Fernandez (2018) abordan también el problema de la lenta certificación de los logros de los estudiantes. A su solución la llaman CredenceLedger, en la que se requieren permisos de acceso, con la idea de hacer extensible a interesados y otras organizaciones la información almacenada en los bloques. Estos autores abogan por una descentralización de los datos. No obstante proponen centralizar la verificación de los

certificados expedidos en una plataforma basada en el uso de la tecnología *blockchain* pero con permisos de acceso. Justifican este modelo ya que puede tender a ser más eficiente en términos de altas tasas de transacciones, bajos costes y bajo consumo de recursos.

Además del problema de lentitud en la consulta de certificados existe el problema de la atribución falsa de los mismos. Dima et al. (2018) pretenden resolver este caso concreto aportando una solución en progreso de reivindicación de la identidad basada en la aplicación de la tecnología *blockchain* pero de acceso privado. Las innovaciones que aportan esta solución, también basada en automatizaciones por *Smart Contracts*, es por un lado la firma múltiple. De esta manera la transacción es válida cuando es firmada por el estudiante y distintas instituciones adheridas a la red. Por otro lado, se puede derogar el diploma expedido, concepto que sobresa por el principio de inmutabilidad.

Un ámbito que también se aplica en el educativo son los dispositivos interconectados, o Internet de las Cosas (Gubbi, Buyya, Marusic, & Palaniswami, 2013; Luigi, Antonio, & Giacomo, 2010). En el aprendizaje ubicuo los dispositivos interconectados tienen un papel fundamental. Otorgan acceso a plataformas educativas, recursos abiertos y a otros artefactos y proceso de aprendizaje desde cualquier lugar. No obstante, Bdiwi et al. (2017) exponen que los sistemas ubicuos presentan problemas de vulnerabilidad. En su trabajo enuncian una solución que implementa la tecnología *blockchain* para preservar los beneficios de seguridad y privacidad en entornos de aprendizaje colaborativos. De la misma manera que otras soluciones, usan *Smart Contracts* para automatizar y mejorar la seguridad de las transacciones. Es un binomio en el que basan su infraestructura de sistema para mejorar los retos de interoperabilidad en educación. Las plataformas educativas están construidas sobre tecnologías que socavan la privacidad de los estudiantes. Es muy fácil hacer el seguimiento, ya no solo en su navegación por Internet, sino por entornos virtuales de aprendizaje en los que se usa *Learning Analytics*. Las leyes de protección de datos personales hacen un intento de otorgar el poder y control de los datos a los propios usuarios. Las leyes pueden ser más efectivas si las herramientas y plataformas educativas se construyen pensando en la privacidad desde el diseño y por defecto. Otorgar el control de los datos a los

estudiantes es lo que Han et al. (2018) pretenden en su solución educativa que implementa la tecnología *blockchain* y de acceso privado, a la vez que proveedores educativos puedan expedir certificados oficiales como prueba de compleción o logro. La extensión de su solución tiene el objetivo de abrazar tanto instituciones educativas formales, como no formales e incluso extraescolares. Los autores de esta propuesta atacan directamente a las universidades como las únicas certificadores oficiales, tildando de sistema anticuado e inadecuado. Abogan por una arquitectura de datos descentralizada y otorgan a los protocolos de consenso todo el peso oficial de cualquier certificación. En esta solución, se vuelve a usar la tecnología *blockchain* para la certificación de títulos, donde los bloques señalan mediante URL (*Uniform Resource Locator*) en qué base de datos se encuentra alojado el documento-título.

Existe una preocupación en cuanto a recolectar datos educativos tanto en contextos formales como no formales o informales. Gong et al. (2019) presentan tres problemas en educación como son la seguridad en datos, en evaluación y en aplicación. Tienen una visión de múltiples desarrollos paralelos que usan la tecnología *blockchain* como solución global a estos problemas. Se reduce la utilidad de la tecnología *blockchain* al hecho de que es una tecnología para seguridad de datos criptográfica basada en almacenamiento distribuido, de evaluación de la seguridad basada en protocolos de consenso y de una seguridad basada en *Smart Contracts*. Su solución guarda en los bloques los enlaces a datos almacenados en una base de datos central.

Las trazas que deja un estudiante a lo largo del tiempo dibujan un contexto complejo. El estudiante participa en distintas instituciones educativas, cada una de ellas con su propia certificación, base de datos local con sus registros de acceso y metadatos, y sus propios procesos de *Learning Analytics*. La solución propuesta por Ocheja, Flanagan y Ogata (Flanagan & Ogata, 2018; Ocheja et al., 2018) tiene como objetivo facilitar la consulta de todas las trazas generadas por los estudiantes a lo largo de su historial de aprendizaje. De esta manera pretenden solucionar la incapacidad de los estándares de interoperabilidad como IMS Caliper o xAPI de unir todas las distintas trazas generadas. Su propuesta usa tecnología *blockchain* para almacenar todas las trazas de los EVA en los que han participado los estudiantes, así como de los almacenes de registros de

aprendizaje en los figuran. A la vez, sitúan a la tecnología *blockchain* como una tecnología segura, en el que los bloques apuntan a la base de datos convencional donde están almacenados los registros de datos y las políticas de accesos. De la misma manera que otras soluciones, usan *Smart Contracts* para automatizar los procesos de acceso y consulta a datos.

La preocupación para proteger la privacidad y datos de los estudiantes en procesos de *Learning Analytics* con tecnología *blockchain* es extensible más allá de los esfuerzos de Ocheja et al. (2018). Farah et al. (2018) proponen con sus sistema que los alumnos tomen el control de sus datos incluso escogiendo el lugar de almacenamiento. Con *Smart Contracts* pretenden asegurar la privacidad de los datos ejecutando de forma automática políticas de acceso contractuales. Un elemento al que hacen referencia es al Reglamento General de Protección de Datos, aspecto que muchos otros autores no tienen en cuenta y que consideran que su solución puede cumplir de alguna forma.

Gilda y Mehrotra (2018) presentan un problema distinto a las certificaciones de logros o trazas en el uso de *Learning Analytics* como es el consentimiento firmado de los padres. Para agilizar la compleción de formularios de consentimiento y permitir un seguimiento del uso de estos, aportan una solución basada en tecnología *blockchain* y de acceso privado donde una serie de *Smart Contracts* permiten que los permisos puedan ser concedidos en cascada, por ejemplo, de la institución educativa a la empresa que conduce las actividades extraescolares. Los autores son conscientes que su propuesta tiene riesgos de seguridad y posible filtrado de datos.

En general, las soluciones propuestas son trabajos en progreso que adoptan el presupuesto de que la tecnología *blockchain* es segura, privada y de bajo coste. Distintos autores la proponen como solución a problemas educativos relacionados con la expedición de certificados, verificación de caminos de aprendizaje, reducción de fraudes en las titulaciones, pasaportes de aprendizaje a lo largo de la vida, gestión de la propiedad intelectual, gestión de datos, compartir recursos educativos mediante *Smart Contracts*, proteger la propiedad intelectual, hacer un seguimiento de las actividades en las que ha participado tanto un profesor como un alumno o incluso compartir registros y resultados de procesos *Learning Analytics* (G. Chen, Xu, Lu, & Chen, 2018; Sun, Wang,

& Wang, 2018; Turcu, Turcu, & Chiuchișan, 2018). No obstante, la arquitectura de la tecnología *blockchain* no está diseñada para almacenar datos, sino punteros a datos, donde los registros académicos, metadatos y certificados se almacenan fuera.

Cabe destacar que algunos autores reportan varios problemas que giran en torno al cumplimiento de leyes como el RGPD y la seguridad de la arquitectura tecnológica *blockchain* (Turcu et al., 2018).

II.8.3.2. RQ2. ¿Qué problemas de seguridad presenta la tecnología *blockchain*?

La seguridad en informática es una temática más acotada que la seguridad de la información. Esta va más allá de la seguridad física de los ordenadores, atiende a defender los sistemas digitales y hacer ilegibles los datos almacenados ante accesos indebidos.

El concepto de seguridad en la tecnología *blockchain* toma una perspectiva distinta. Se refiere a evitar modificaciones en sus bloques. En consecuencia, tanto en informática como en la unión de bloques de transacciones mediante criptografía, la seguridad trata de evitar ataques maliciosos para acceder a la información (Landwehr, 2018).

La tecnología *blockchain* aporta confianza en las transacciones entre usuarios. Esta confianza se basa en la transparencia y validación por consenso, que en resumen su seguridad radica en el principio que ninguna entidad debe poseer más del 51% o de lo contrario adquiriría en control absoluto de la red (Gervais et al., 2016).

Para validar las transacciones y añadir bloques en la cadena se pueden usar distintos protocolos de consenso. El protocolo de consenso más conocido es el *Proof of Work* (PoW, en español Prueba de Trabajo), el cual es usado en el 90% de las criptomonedas existentes (Gervais et al., 2016). No obstante, estos autores destacan que este protocolo presenta vulnerabilidades que hacen que las implementaciones de la tecnología *blockchain* sean inseguras y necesarias de revisión profunda. Existen otras propuestas de protocolos de consenso como PoS (*Roof of Stake*), PoB (*Proof of Burn*) o PoC (*Proof of Capacity*). Gervais et al. (2016) presentan el ataque llamado *selfish mining* (en castellano, minería egoísta o maliciosa). Este ataque consiste en generar en muy poco

tiempo distintas bifurcaciones de cadena para conseguir realizar un doble gasto. Así mismo se basa en el principio de PoW de la cadena más larga como cadena válida, de manera que en el ataque se consigue validar una cadena alterada. En sus conclusiones, los autores determinan que la aplicación de la tecnología *blockchain* en *Bitcoin* ofrece más seguridad que la aplicada en *Ethereum*.

La presente investigación identifica la poca privacidad de las infraestructuras protocolarias de Internet, motivo por el cual es posible capturar datos con facilidad e incluso alterarlos. Apostolaki et al. (2017) aprovechan el protocolo de enrutamiento de Internet (BGP, del inglés *Border Gateway Protocol*) para realizar ataques consistentes en aislar partes de la red de nodos (*partitioning attacks*) o retardar la propagación de bloques (*delay attacks*). Estos ataques pueden hacer perder mucho dinero e incluso realizar doble gasto. Lo destacable del trabajo de Apostolaki et al. (2017) es que demuestran que *Bitcoin* es centralizado. Esta centralización es a nivel de enrutamiento, puesto que el 32% del tráfico de la Red se concentra en Hurricane Electric (Hurricane Electric, 2019). Junto a Level3 y Telianet suman el 60% de todas las conexiones posibles de *Bitcoin*. Hurricane Electric es considerado un AS (*Autonomous System*, en castellano Sistema Automático), cuya capacidad de concentración superior a cualquier nodo de la red puede efectuar estos ataques de particionado o retardado. Por consiguiente, otra de las debilidades de la aplicación de la tecnología *blockchain* reside en las vulnerabilidades que presenta un AS con alta concentración de las conexiones de sus nodos.

En otro tipo de ataque, como el ataque 51%, una solución que use tecnología *blockchain* de acceso privado como *Ethereum* necesita tener suficiente masa crítica para evitar que se alteren las inserciones de las transacciones. El protocolo de consenso PoW funciona muy bien con grandes redes de usuarios. Cuando se traslada la prueba de consenso a compañías privadas, puede que en este contexto localizado la masa crítica de usuarios de sus servicios no sea suficiente para asegurar la no manipulación de los bloques de transacciones (*tamper-proof*). Apostolaki et al. (2017) analizan la relación entre seguridad y rendimiento de los protocolos de consenso en un trabajo que pone en entredicho la capacidad de asegurar que los bloques no sean alterados.

Los usos de la tecnología blockchain redefinen el concepto de confianza utilizando criptografía y protocolos de consenso para asegurar niveles de privacidad, seguridad, anonimato e integridad de los datos sin la necesidad de una entidad central (F. Dai, Shi, Meng, Wei, & Ye, 2017). No obstante, estos autores ponen de manifiesto algunas limitaciones técnicas, entre ellas el restrictivo tamaño de los bloques, el mecanismo de almacenamiento distribuido que proporciona copias de todos los datos a posibles atacantes y la posibilidad de conseguir un 51% de poder en los protocolos de consenso. De la misma manera que Drescher (2017), apunta a una debilidad base en cuanto a la posible pérdida de la clave privada por parte de un usuario. La solución de los gestores de claves privadas, como Coinbase o Binance, ponen en peligro la privacidad de los usuarios al estar abiertos a ataques convencionales de tecnologías web.

Drescher (2017) ofrece una visión general de las limitaciones técnicas a las que hacen referencia Dai et al. (2017), algunas de ellas centradas en la privacidad y seguridad del sistema. Drescher expone que las soluciones basadas en tecnología *blockchain* tienen una falta de privacidad al estar los datos copiados en cada uno de los usuarios de la red, una seguridad tan débil como la pérdida de la clave privada de la encriptación asimétrica, una centralidad oculta en aquellos que tienen suficiente capacidad de cómputo para llegar al 51% y una alta probabilidad de conseguir un 51% por pocos nodos en aquellas redes con pocos participantes. En cuestiones no técnicas también apunta a algunas limitaciones, como la falta de cumplimiento en cuestiones legales, contexto que ya se vivió con la adopción de Internet en los años 90. Las soluciones a estas limitaciones pueden acarrear cambios drásticos en los fundamentos, del mismo modo que deberían hacerse en internet. Drescher expone que las soluciones no tecnológicas pasan por una comprensión de la tecnología por parte de la sociedad y una adaptación del nuevo contexto tecnológico a las leyes, o nuevas leyes.

Otros autores como Conti et al. (2018) analizan problemas de seguridad y privacidad en *Bitcoin*, o lo que es lo mismo, en la aplicación de la tecnología *blockchain*. Así, estos autores exponen como *Bitcoin*:

- No requiere una entidad central de gestión, no obstante autores como Apostolaki et al. (2017) demuestran que el 60% del enrutamiento de las conexiones de la red de nodos se centralizan en tres supernodos.
- Aporta anonimato y transparencia, aunque los autores señalan que la seudonimización que ofrecen las soluciones que usan tecnología *blockchain* puede romperse y terminar identificando a los propietarios de las transacciones.
- Ofrece reducción de tasas, no obstante, los autores apuntan a que esta reducción solo es beneficiosa para aquellos usuarios que mueven grandes cantidades de dinero, no siendo así para pequeñas transacciones.
- Es seguro, aunque Conti et al. (2018) exponen como distintos vectores de ataque (*selfish mining, block discarding attack, Pool Hopping attack, Bribery attack, Sybil attack* o *Eclipse attack*) ponen en entredicho la seguridad y permite realizar doble gasto. Estos autores también resaltan los peligros del uso de la tecnología por parte de clientes, los cuales quedan expuestos a robos de su cartera digital mediante *hacking, malware* o mal uso de esta.

Es importante notar como los autores indican que la privacidad no es una propiedad considerada en el diseño inicial de *Bitcoin*, y por ende ni en el diseño de la tecnología *blockchain*, tecnología en la que muchos investigadores destacan carencias y peligros.

Los mismos ataques a soluciones basadas en tecnología *blockchain* analizados por Conti et al. (2018) también los analizan Soni y Maheshwari (2018), con la diferencia que ellos además analizan los protocolos surgidos como defensa de los ataques. En su trabajo, Soni y Maheshwari presentan distintos ataques como *eclipse attack, Sybil attack, >51% attack* y *selfish mining*. Por la literatura seleccionada se infiere que estos ataques son los más comunes, y según Soni y Maheshwari aún restan vulnerabilidades por descubrir. Aseguran que los ataques aún requieren un estudio en profundidad y, a pesar de los avances, hasta ahora no se ha encontrado una solución satisfactoria.

Conocer los problemas de seguridad que presentan las implementaciones de *blockchain* más usadas, como *Bitcoin, Ethereum* o *Hyperdelger*, es una forma de comprender que no hay una única aplicación de la tecnología *blockchain* y una única manera de abordar

la seguridad. Este hecho también explica, como apuntan Conti et al. (2018), que la arquitectura de la tecnología *blockchain* carece de privacidad desde el diseño. Moubarak, Filiol y Chamoun (2018) presentan un su trabajo distintos ataques relacionados con la seguridad en las transacciones, ataques *spam*, *smart contracts* maliciosos, ataques DDoS (*Distributed Denial of Service*) o ataques *timejacking*. Los autores, y tras analizar distintas vulnerabilidades de seguridad y ataques de las aplicaciones de la tecnología *blockchain* como *Bitcoin*, *Ethereum* o *Hyperledger*, no recomiendan el uso de ninguna de ellas. Advierten que el eslabón más débil de la cadena de bloques es la capacidad de salvaguardar la clave privada del usuario.

Autores como Joshi et al. (2018) y Bacon et al. (2018) analizan directamente la arquitectura de la tecnología *blockchain*, su diferentes algoritmos de consenso así como los retos y oportunidades desde la perspectiva de la seguridad y privacidad. Bacon et al. (2018) apuntan ataques a vulnerabilidades como el *>51% attack*, *Sybil attack*, *DDoS attacks* o posibles errores humanos en *Smart Contracts*. Los autores Joshi et al. (2018) hacen un recorrido de las oportunidades de aplicar la arquitectura de bloques en contextos como las finanzas, la salud, las aplicaciones móviles, los automóviles o incluso defensa ante ciberataques. Asimismo, también expone los retos a los que se enfrentan las posibles soluciones que implementan la tecnología *blockchain* como problemas en escalabilidad, filtrado de la privacidad, ataque *selfish mining*, información personal identificable o seguridad. No obstante, los autores creen que el uso de la tecnología *blockchain* será extendido a pesar de las actuales limitaciones y posibilidades de crear aplicaciones innovadoras.

Las vulnerabilidades en los *Smarts Contracts* de *Ethereum*, y por consiguiente a considerar en cualquier otra aplicación de la tecnología *blockchain*, son catalogadas por Mense y Fltascher (2018) en una lista taxonómica en la que se agrupan por alta severidad, severidad no crítica y alerta. Se destaca que hay más vulnerabilidades de alta severidad. De las 22 vulnerabilidades presentadas, 15 (68%) son de alta severidad, 8 (36%) de severidad no crítica y 7 (31%) de alerta. La suma de las vulnerabilidades de severidad no crítica y alerta igualan a las de severidad crítica.

II.8.3.3. RQ3. ¿La tecnología blockchain cumple con el RGPD?

La creación de la tecnología *blockchain* por parte de Nakamoto (2008) permite redefinir con sus usos el concepto de confianza utilizando criptografía y protocolos de consenso. No obstante, la privacidad no es una característica presente en su diseño según aseveran Conti et al. (2018), hecho que puede hacer incumplir el RGPD. En la revisión sistemática de la literatura se encuentran autores de conclusiones contrapuestas, que dan a entender que puede usarse y aplicarse en condiciones tecnológicas y de políticas muy concretas que puedan cumplir con el RGPD.

Autores como Al-Zaben et al. (2019) exponen el estado distópico que se presenta en el capítulo II.2 Internet insegura y *clickstream*. Muestran como los incidentes sobre vigilancia y brechas en la privacidad de los usuarios permiten almacenar grandes cantidades de información personal identificable, explotada por compañías para análisis y predicción de mercados basados en datos. Los autores consideran que no poder hacer un seguimiento de los datos es un problema para los usuarios. Proponen una solución basada en tecnología *blockchain* que permite hacer este seguimiento y que además cumple con el RGPD. Para conseguirlo proponen una tecnología híbrida de base de datos *off-chain* (fuera de los bloques de transacciones) para almacenar los datos y una implementación de la tecnología *blockchain* para mantener un ciclo de vida confiable de datos personales identificables. Estos autores demuestran que una solución que solo implemente tecnología *blockchain* no puede cumplir con el RGPD y necesita de una base de datos local para, al menos, cumplir con los derechos de olvido y borrado de datos personales. No obstante, los autores aseguran que su solución híbrida cumple con el RGPD al disponer de una arquitectura separada en tres roles: usuario, controlador y procesador, y resolver dos de los derechos que el RGPD dota a los usuarios, el derecho al olvido y el derecho al borrado de los datos. No hay referencia en cuanto a procesos de seguridad y privacidad en las bases de datos locales, ni quien es el delegado de protección de datos que se haga cargo de los datos almacenados en los bloques de transacciones definidos en la tecnología *blockchain*.

Aunque Al-Zaben et al. (2019) expongan una solución híbrida aparentemente sencilla en el cumplimiento del RGPD, Millard (2018) pone de manifiesto que los procesos en línea

están más regulados que los fuera de línea, hecho que dificulta sobremanera regular y hacer cumplir las regulaciones a las tecnologías emergentes. Los autores se preguntan “si es posible construir y desplegar plataformas que implementen tecnología *blockchain* compatibles con el RGPD” en la medida que implican el tratamiento de datos personales. Para dar respuesta hace referencia a personas destacadas dentro del proceso de confección del RGPD, como Jan Philip Albrecht, cuyos comentarios aseguran “que probablemente cualquier implementación de la tecnología *blockchain* no pueda utilizarse para procesar datos personales”.

Millard (2018) recorre las características de la tecnología de bloques, los requisitos del RGPD y da una respuesta definitiva a su pregunta con un “quizás” y un “depende”. Cualquier implementación de la tecnología *blockchain* en su estado original no cumple con el RGPD. No obstante, Millard argumenta que, si se usa como entorno privado y con permisos mediante un consenso por autoridad, entonces podría ser que cumpliera con el RGPD. Asegura que aún queda mucho trabajo por hacer en cuanto a legalidad y no solo en cuestiones de RGPD. Se destaca que la aproximación de Millard rompe con los principios de transparencia, inmutabilidad y autoridad distribuida. Surge la pregunta si la tecnología *blockchain*, privada de sus principios fundamentales, se convierte en una simple lista de datos ordenada por tiempo de inserción (ver sección II.7.4 *Blockchain* y DLT).

Otros autores como Pagallo et al. (2018) son más contundentes en sus afirmaciones y aseguran que no cumple con el RGPD, pero llegará el momento en que puede ser factible. En su trabajo, muy centrado en el Artículo 17 en relación con el derecho de supresión, distinguen aquellos usos de *blockchain* pre RGPD y post RGPD, siendo los creados después de la aparición del RGPD los que pueden cumplir con este. Discurren entre las similitudes de origen entre las redes P2P y redes que dan soporte a la tecnología *blockchain*, en cuanto sus usos pueden ser ilegales.

Pagallo et al. (2018) consideran tres posibles soluciones para que un sistema que implemente la tecnología *blockchain* cumpla con las normas del RGPD o incluso para impedir que tales sistemas procesen datos personales. No obstante, por exposiciones de los autores, todas presentan sus propios inconvenientes, por ello proponen:

- Guardar *hashes* en los bloques de transacciones como enlaces a datos almacenados fuera del mismo (*hashing-out*), sin embargo, puede considerarse una traición al principio de descentralización de las cadenas de bloques, ya que un cierto grado de control de los datos permanece en manos de un único partido centralizado.
- Destruir las llaves de encriptado, de manera que no puedan descifrarse los datos almacenados, sin embargo, esto presenta problemas futuros en cuanto a evolución de la tecnología y capacidad superior de revertir la encriptación y finalmente provocar una identificación.
- Definir *hashes* camaleónicos de manera que mediante una puerta trasera se puedan modificar bloques a conveniencia, sin embargo, los “*hashes* camaleónicos no pueden eliminar las antiguas copias de la *blockchain*, que aún contienen la información, sin tener en cuenta que los mineros también tienen la discreción de aceptar las nuevas cadenas o quedarse con las antiguas”.

Los autores muestran algunas soluciones que hasta ahora presentan sus propios inconvenientes. Esto los lleva a una paradoja y a preguntarse “¿se debe admitir que no existe una solución única en este contexto, salvo la prohibición de la tecnología?”. Su conclusión es que los principios de la tecnología *blockchain* deben ser traicionados para poder cumplir con el RGPD, puesto que la privacidad no está implementada en su diseño ni por defecto como apunta el Artículo 25. Además, se requiere una autoridad central que regule los datos añadidos en los bloques, diferenciando así los usos de la tecnología *blockchain* pre RGPD con serios problemas legales y los *post* RGPD que de alguna forma conseguirán cumplir con este, aunque los autores no saben cómo.

Autores como Bacon et al. (2018) apuntan a que la aplicación del RGPD plantea preguntas difíciles. Notan que las soluciones que implementan tecnología *blockchain* están afectadas por el RGPD y que la transferencia internacional de datos dificulta la legalidad impactando a responsables y encargados del tratamiento de distintos países. Los autores recorren los distintos aspectos del RGPD más importantes como:

- La transferencia internacional de datos, puesto que el reglamento afecta a responsables y encargados del tratamiento de distintos países.

- El concepto de dato personal, que es toda información que pueda identificar directa o indirectamente a una persona. De esta manera los sistemas que usan tecnología *blockchain* deben cumplir el RGPD ya que los metadatos de los bloques llevan identificadores considerados datos personales. Exponen vulnerabilidades y posibilidad de identificación al mezclar datos de los intermediarios, como *CoinBase*, con los almacenados en los bloques de transacciones. Proponen la solución de almacenar datos personales en una base de datos *off-chain*, solución propuesta por Al-Zaben et al. (2019) y Pagallo et al. (2018) con los inconvenientes apuntados por estos autores.

Bacon et al. (2018) dibujan un contexto complicado donde los responsables y encargados de datos personales se asignan según el propósito que tengan al usar los datos personales. Liberan a los desarrolladores de cumplir con el RGPD, sin embargo, juegan un papel muy importante en la gobernanza de los datos y del sistema que los permite compartir. Los autores concluyen que:

- El uso de tecnología *blockchain* de acceso privado aumentará.
- No está claro que el uso de la tecnología *blockchain* de acceso privado y, por tanto, centralizado, ofrezca ventajas significativas sobre las soluciones de bases de datos existentes, especialmente cuando los sistemas gestores de bases de datos pueden realizar muchas de las mismas funciones.
- Alcanzar la legalidad es más difícil en las plataformas abiertas y distribuidas que deliberadamente carecen de un administrador central con control sobre el libro mayor.
- Si se utiliza una plataforma que implemente la tecnología *blockchain* para procesar datos personales, entonces los usuarios, nodos y mineros pueden ser responsables de datos, encargados de datos o incluso ambos. En el caso de las plataformas abiertas y distribuidas, incluso si todas las partes implicadas fueran consideradas controladores conjuntos, no está claro cómo cumplirían sus obligaciones.

II.8.3.4. RQ4. ¿Qué puede resolver la tecnología blockchain en relación con el problema?

Se define el problema de la investigación como la fragilidad en cuanto a la protección de la privacidad de datos e identidad de los estudiantes en el uso de *Learning Analytics*. Esto se traduce en que los datos personales almacenados y posteriormente tratados o compartidos por terceros se vean tanto asegurada su privacidad como su seguridad ante vulnerabilidades del sistema.

Se espera que las implementaciones de la tecnología *blockchain* resuelvan el problema en parte o completamente. La palabra privacidad, tal y como se ha considerado utilizar en esta investigación, abraza cuestiones como intimidad, anonimato, seudonimización o confidencialidad de datos. Para que las implementaciones de la tecnología *blockchain* sean soluciones reales deben abordarse todos estos conceptos y también cumplir con la legalidad.

En cuestiones de identidad, y preservar la misma en procesos de *Learning Analytics*, se pueden considerar las palabras de Dai et al. (2017) en sus afirmaciones:

Blockchain adopta un mecanismo de cifrado asimétrico que permite a los usuarios cifrar los datos con su propia clave privada. Además, el valor *hash* de la clave pública de un usuario se calcula y actúa como indicador de ID del usuario. Por un lado, el valor *hash* no tiene relación con la identidad real del usuario, por lo que se mantiene segura la información de privacidad personal del usuario. Por otro lado, el proceso de cálculo del valor *hash* es invertible, lo que significa que un adversario no puede calcular la clave pública de un usuario a partir de la dirección pública del usuario, y es imposible calcular la clave privada a partir de la clave pública. Por lo tanto, *blockchain* logra el objetivo de preservar el anonimato y la privacidad de los usuarios (pp. 2).

Se entiende de las palabras de Dai et al. (2017) que los estudiantes que operen en soluciones que implementan la tecnología *blockchain* preservarán su anonimato. Como mínimo, mientras la clave privada utilizada para generar el hash final esté en buen

recaudo, el alumno tiene opciones de permanecer en el anonimato. Este precepto define el punto más débil, puesto que un alumno menor de edad necesitará que le gestionen sus claves privadas. En tal caso la seguridad incidirá en el gestor de las claves, que deberá utilizar métodos convencionales de almacenamiento incluso accesibles vía web (Al-Zaben et al., 2019), con las consecuentes vulnerabilidades de la arquitectura Internet-Web (ver II.2 Internet insegura y *clickstream*). Asimismo, las distintas vulnerabilidades (Conti et al., 2018; Joshi et al., 2018) ponen en entredicho la capacidad de las aplicaciones actuales de la tecnología *blockchain* de preservar el anonimato y confidencialidad de los datos almacenados en este, o de preservar la integridad de los mismos.

Los responsables de la gestión de claves y datos personales de los estudiantes deben cumplir la legalidad, en concreto el RGPD. Por consiguiente, una solución que use tecnología *blockchain de acceso público* y sin permisos no puede o difícilmente puede cumplir con el RGPD (Al-Zaben et al., 2019; Bacon et al., 2018; Pagallo et al., 2018). Es necesario crear sistemas híbridos, de otra forma es muy complejo que una solución que solo implemente la tecnología *blockchain* cumpla con la legalidad y sea funcional desde la perspectiva técnica como capacidad de espacio o transferencia entre nodos.

En este punto de la revisión de la literatura se plantea la siguiente pregunta que asimismo dará respuesta a la RQ4, ¿si una solución implementada con tecnología *blockchain* no puede almacenar datos personales y debe hacerlo como se ha hecho siempre en bases de datos locales, qué aspectos resuelve en cuanto a privacidad y seguridad fuera de *blockchain*? Es importante resolver esta pregunta puesto que los datos personales de los estudiantes están almacenados precisamente en base de datos *off-chain*.

Gilda y Mehrotra (2018) y Flanagan y Ogata (2018), así como los autores presentados en la pregunta RQ1. *¿Qué soluciones se han aportado en el campo de estudio?* aportan la respuesta definitiva. *Blockchain* es una tecnología que permite certificar sucesos en un entorno de confianza por consenso y firmar autoría mediante criptografía. Sin embargo, a pesar de que se considere una base de datos, no es una tecnología diseñada para almacenar gran cantidad de datos. Por consiguiente, se remite a las soluciones de los

autores presentados a lo largo de la revisión, permite una gestión supuestamente segura y privada de los datos personales almacenados en bases de datos fuera de sus bloques. Por ejemplo, resuelve situaciones de gestión de datos en procesos de interoperabilidad o de trazabilidad de los datos entre las distintas entidades implicadas en procesos de *Learning Analytics*.

II.8.4. Análisis resumen de las soluciones propuestas

A lo largo de las distintas respuestas a las preguntas de investigación se ha revisado cada uno de los 30 trabajos seleccionados y relacionados con dominios de los ámbitos “educación” y “*blockchain*”. Se concluye que es imperativo abordar el problema fuera de la cadena de bloques que define la tecnología *blockchain*, puesto que añade complejidad y no lo soluciona. La investigación se centra en el contexto de almacenamiento fuera de la cadena para entregar a los alumnos un nivel adecuado de seguridad y confidencialidad de datos, y anonimato de su identidad.

En la Tabla 19, y a modo de resumen, se comentan dichos trabajos junto a las conclusiones más trascendentes de cada pregunta de investigación.

Tabla 19 Resumen de los resultados de la revisión sistemática

Pregunta	Trabajos
RQ1	Distintos autores proponen implementar la tecnología <i>blockchain</i> para solucionar problemas educativos relacionados con la expedición de certificados, verificación de caminos de aprendizaje, reducción de fraudes en las titulaciones, pasaportes de aprendizaje a lo largo de la vida, gestión de la propiedad intelectual, gestión de datos, compartir recursos educativos mediante <i>Smart Contracts</i> , proteger la propiedad intelectual, hacer un seguimiento de las actividades en las que ha participado tanto un profesor como un estudiante o incluso compartir registros y resultados de procesos <i>Learning Analytics</i>
RQ2	<i>Blockchain</i> es una tecnología que por su diseño presenta una serie de vulnerabilidades de seguridad y de privacidad que en su implementación son explotables con ataques como el <i>>51% attack</i> o <i>Sybil attack</i> , entre otros. Según se utilice la tecnología <i>blockchain</i> para crear criptomonedas o soluciones para

distintos contextos, surgen otras vulnerabilidades como errores en humanos en el desarrollo *Smart Contracts* en *Ethereum* o *delay attacks* en *Bitcoin*

- RQ3 El diseño inicial de la tecnología *blockchain* dificulta enormemente su cumplimiento con el RGPD. Romper con el principio de autoridad descentralizada es la solución generalizada para que las soluciones que implementen la tecnología *blockchain* cumplan con el RGPD. No obstante, surge la duda si en ciertos casos es necesario utilizar una implementación de la tecnología *blockchain* puesto que soluciones de bases de datos tradicionales cumplen con los mismos objetivos.
- RQ4 Resuelve situaciones de gestión de datos en procesos de interoperabilidad entre herramientas o de trazabilidad de los datos entre los distintos responsables y encargados de los datos personales en el uso de *Learning Analytics*. No obstante, carece de la capacidad de resolver cuestiones de privacidad y seguridad que pueden encontrarse en bases de datos, como pueden ser excepciones como el ejercicio del derecho de anonimato en un entorno virtual de aprendizaje.
-

De estos se extrae que la arquitectura de la tecnología *blockchain* no está diseñada para almacenar datos más allá de los transaccionales, puesto que supone descargar una copia del libro mayor en todos los usuarios de la red. Guardar todos los datos en los bloques de transacciones provoca una congestión de la propia red, del almacenamiento de los propios usuarios y cuestiones adicionales sensibles en cuanto a protección de datos y legalidad vigente. Para agilizar la conectividad entre los distintos nodos se utilizan estructuras digitales como el árbol Merkel (Duan et al., 2018; Gervais et al., 2016; Moubarak et al., 2018), donde los bloques funcionan como punteros a datos almacenados fuera de su red.

Se extrae también que usar la tecnología *blockchain* para almacenar enlaces a datos en sistemas gestores de bases de datos es la aproximación menos arriesgada y más cercana a la legalidad (Mense & Flatscher, 2018; Pagallo et al., 2018). Es de esta arquitectura de enlaces a datos externos de la que parten todas las soluciones educativas encontradas, así como otras que quieren cumplir con la legalidad vigente.

II.8.4.1. Soluciones educativas

Por lo general, el uso de la tecnología *blockchain* se relaciona con “garantizar certificación”. Sin embargo, las promesas de los interesados en esta tecnología van a veces más allá de una simple certificación temporal de los sucesos educativos (Wilkinson & Lowry, 2014). Su aplicación en educación pretende resolver situaciones de:

- Toma de mejores decisiones con datos.
- Certificación de logros de aprendizaje
- Interconexión de dispositivos.
- Seguimiento de trazas de aprendizaje entre proveedores.

Bore et al. (2017) proponen utilizar la tecnología *blockchain* para tomar mejores decisiones basadas en datos. Con dicho uso aseguran que todos los datos de cualquier institución educativa de un país estén centralizados a tiempo real y se asegure su inmutabilidad. Se ofrece un marco de ingeniería de datos y *Big Data* en el que agentes gubernamentales y de distintas organizaciones puedan consultarlos para tomar decisiones en base a datos.

En el marco de la certificación efectiva de conocimiento, habilidades y logros de los estudiantes, distintos autores presentan soluciones donde la implementación de la tecnología *blockchain* permite enlazar con autoría a los certificados expedidos. Duan, Zhong y Liu (2018) exponen en su trabajo cómo el uso de la tecnología *blockchain* puede ayudar a sobrepasar el problema de la falta de efectividad en la verificación; Xu et al. (2017) buscan en su trabajo un sistema que permita maximizar la eficiencia de búsquedas a la par que asegurar la privacidad de los datos de los alumnos en la gestión de certificaciones; y Turkanović et al. (2018) proponen el EduCTX, una solución de acceso global y más atomizada en el contexto de las certificaciones en el marco de los créditos ECTS.

En este proceso de certificación Han et al. (2018) proponen un sistema que implemente la tecnología *blockchain* para otorgar el control de los datos a los propios alumnos. No son los únicos, puesto que Farah et al. (2018) proponen con su sistema que los estudiantes tomen el control de sus datos e incluso escoger el lugar de almacenamiento.

Introducen los *Smart Contracts* como tecnología adicional para ejecutar de forma automática políticas de acceso contractuales. En la misma dirección Ocheja et al. (Flanagan & Ogata, 2018; Ocheja et al., 2018) procuran facilitar la consulta de todas las trazas generadas por los estudiantes en los entornos virtuales de aprendizaje a lo largo de su historial de aprendizaje.

Otros autores como Gilda y Mehrotra (2018) se preocupan por la compleción de formularios de consentimiento y el seguimiento de estos, que a la par que Farah et al. (2018) usan *Smart Contracts* para automatizar políticas.

La ubicuidad de los dispositivos en educación no se escapa del uso de la tecnología *blockchain* para solucionar aspectos concretos al respecto. En este contexto, atendiendo a las posibles vulnerabilidades en la interconexión de dispositivos, Bdiwi et al. (2017) exponen una solución desarrollada con la tecnología *blockchain* para preservar los beneficios de seguridad y privacidad en entornos de aprendizaje colaborativos. Gong et al. (2019) presentan preocupaciones en la seguridad de datos y una solución basada en múltiples sistemas *blockchain* paralelos, donde cada uno de los nodos enlazan a datos *off-chain*.

II.8.4.2. Seguridad transaccional

Las propuestas de soluciones a problemas educativos se basan en los principios de seguridad criptográfica. No obstante, la tecnología *blockchain* está construida sobre una infraestructura de red y se acompaña de tecnologías de fácil acceso como la Web, que socavan su seguridad. Por consiguiente, a los ataques a las implementaciones de la tecnología *blockchain* se le suman los propios ataques de la Web.

Es importante notar como algunos autores indican que la privacidad (Conti et al., 2018), al igual que sucede con las tecnologías Internet y Web (Al-Zaben et al., 2019), no es una propiedad de *Bitcoin* (basado en tecnología *blockchain*) considerada en su diseño inicial. Muchos investigadores han presentado carencias, peligros en este campo y ataques que debilitan las redes de nodos que dan soporte a la tecnología *blockchain*, socavan la seguridad criptográfica, y por ende la privacidad hasta el punto de desvelar las identidades de los usuarios (Conti et al., 2018). Lo más preocupante es que aún faltan

soluciones efectivas y definitivas a las distintas vulnerabilidades (Soni & Maheshwari, 2018). Por consiguiente, las distintas vulnerabilidades detectadas ponen de manifiesto una debilidad permanente tanto en sistemas públicos y sin permiso, pero sobretodo en sistemas privados y con permiso donde la verificación por consenso es centralizada. Y el número de vulnerabilidades no son pocas, tal y como apuntan distintos autores:

- *Selfish mining, partitioning attacks* (Gervais et al., 2016).
- *Delay attacks* (Apostolaki et al., 2017).
- >51% (Drescher, 2017).
- *Block discarding attack, Pool Hopping attack, Bribery attack, Sybil attack o Eclipse attack* (Bacon et al., 2018; Conti et al., 2018; Joshi et al., 2018; Soni & Maheshwari, 2018).
- Ataques *spam, Smart Contracts* maliciosos, ataques DDoS o ataques *timejacking* (Moubarak et al., 2018).
- Otras vulnerabilidades de alta severidad detectadas por (Mense & Flatscher, 2018).

Algunas criptomonedas se han visto afectadas por ataques, provocando un gran conflicto de intereses, puesto que los usuarios requieren seguir minando para conseguir ingresos. No obstante, las soluciones pasan por regenerar las cadenas de bloques en el punto del ataque, aspecto que conlleva mucho tiempo y costes de computación, o saltar a otra copia de la cadena de bloques. Esto provoca un gran inconveniente en soluciones educativas, puesto que puede suponer dejar una copia con los datos comprometidos. Se añade que la idea de regeneración rompe con el principio de inmutabilidad presentado por la tecnología *blockchain*, puesto que un acuerdo entre los nodos que aúna un 51% del poder computacional de la red pueden alterar los bloques en cualquier punto deseado. La inmutabilidad que se define en la especificación de la arquitectura de la tecnología *blockchain* es débil, puesto que un retroceso a un estado anterior es posible (Kharif & Marsh, 2019).

II.8.4.3. Legalidad

La implementación de la tecnología *blockchain* creada por Nakamoto (2008) redefine el concepto de confianza utilizando criptografía y protocolos de consenso descentralizado. No obstante, la privacidad no es una característica presente en su diseño (Conti et al., 2018). Esta carencia de privacidad por diseño y por defecto es una cuestión necesaria que el RGPD subraya en el punto 78:

La protección de los derechos y libertades de las personas físicas con respecto al tratamiento de datos personales exige la adopción de medidas técnicas y organizativas apropiadas con el fin de garantizar el cumplimiento de los requisitos del presente Reglamento. A fin de poder demostrar la conformidad con el presente Reglamento, el responsable del tratamiento debe adoptar políticas internas y aplicar medidas que cumplan en particular los principios de protección de datos desde el diseño y por defecto. [...]

Con este solo punto del RGPD se dividen las soluciones desarrolladas con tecnología *blockchain* antes y después de su promulgación (Pagallo et al., 2018). RGPD enmarca un complejo entramado de roles para asegurar una adecuada confidencialidad de datos, como son los propietarios de los datos, los responsables del tratamiento de los datos, los encargados del tratamiento de los datos y el delegado de protección de datos. Autores como Al-Zaben et al. (2019) exponen una solución que contempla a los responsables y a los encargados, aunque como expone Millard (2018), las leyes en contextos en línea son más exigentes, o que existe más regulación, y la solución de Al-Zaben et al. puede caer en incumplimiento.

Es el entramado legal que hace considerar a Millard (2018) que un sistema que usa tecnología *blockchain* quizás pueda en dependencia de cómo esté diseñada su arquitectura cumplir con el RGPD. Bacon et al. (2018) apuntan a que la aplicación del RGPD plantea preguntas difíciles, puesto que el contexto es complicado y la arquitectura de diseño inicial de la tecnología *blockchain* en cuanto a transparencia y consenso

descentralizado debe reformularse. Es un efecto directo del Artículo 17 en relación con el derecho a la supresión (u olvido) de los datos recolectados y tratados.

La conclusión de todos los autores es que las soluciones implementadas con tecnología *blockchain* actuales que pueden cumplir con el RGPD están en progreso de revisión, a expensas de la evolución de las leyes y concreción de conceptos ambiguos.

II.8.4.4. Fragilidad

La presente investigación manifiesta una fuerte fragilidad en la privacidad, confidencialidad y seguridad durante el tratamiento de los datos educativos en procesos de *Learning Analytics*, y dentro de los EVA. Las respuestas a las preguntas de la SLR demuestran una fuerte inestabilidad e incapacidad de las soluciones desarrolladas con tecnología *blockchain* para dar soluciones robustas y sostenibles. Las razones son múltiples y relacionadas con una inmadurez, inseguridad e ilegalidad tecnológica. Su inmadurez, tal y como apunta Gardner (Litan & Leow, 2019), crea unas sobre expectativas en la resolución de muchos problemas, que en el ámbito educativo hace que aún deba esperarse un cierto tiempo de normalización y resultados sólidos para considerarse una tecnología efectiva, puesto que los datos de los estudiantes siguen almacenándose fuera de la cadena. Las distintas vulnerabilidades (ver II.8.4.2 Seguridad transaccional), vectores de ataques y verdadera centralización en la conectividad, confieren una desconfianza y unos peligros que aconsejan una privatización que se aleja de sus principios fundamentales como descentralización e inmutabilidad, y se aproxima al uso de bases de datos relacionales. Y en cuestiones de ilegalidad, solo los principios de inmutabilidad y datos distribuidos plantean serios problemas de compatibilidad con las leyes, quedándose a las expectativas de la evolución tanto tecnológica como reglamentaria.

Se concluye por todo lo expuesto en la SLR, que las soluciones implementadas con tecnología blockchain podrían llegar a resolver cuestiones de interoperabilidad e incluso de certificación, pero en ningún caso el problema de la fragilidad expuesto en la investigación. La fragilidad debe resolverse a nivel de sistema gestor de bases de datos, tecnología utilizada en los EVA en los que incurren procesos de Learning Analytics.

II.8.5. Amenazas a la validez de esta revisión de la literatura

La presente revisión sistemática de la literatura, junto a sus estrategias de mapeo, pretende comprender el campo científico en relación con educación y *blockchain*. No obstante, en una revisión sistemática de la literatura existen ciertas amenazas a la validez e incluso limitaciones incontrolables en los trabajos seleccionados (Neiva et al., 2016). Para evitar posibles desviaciones, sesgos o inclinaciones de los autores, e incluso del propio autor de la presente investigación, se han aplicado una serie de procedimientos como la lista de control de calidad ilustrada en la Tabla 13 (Kitchenham & Charters, 2007).

Las limitaciones también vienen sucedidas por los orígenes seleccionados para la extracción de los trabajos. En la investigación se han utilizado las bases de datos digitales ACM Library, Springer Link, Web of Science e IEEE Xplorer, así como otras fuentes como Google Scholar como literatura gris. A excepción de la última selección, todas las fuentes de datos han sido seleccionadas a razón de su impacto en contexto científico que asegura la calidad de los trabajos indexados.

Poder reproducir la investigación es importante en el ámbito científico, tanto para exponer transparencia como para filtrar sesgos en procedimientos y razonamientos. La última medida consiste en la publicación de la base de datos final, con todas las fases de la revisión sistemática de la literatura, realizada con Microsoft Excel (Amo, 2018).

III. Marco empírico

Como se ha demostrado previamente, las soluciones educativas desarrolladas con tecnología *blockchain* presentan un estado *post* promesas exageradas aún en proceso de madurez (Litan & Leow, 2019), donde en general:

- Se añade una capa más de complejidad.
- Se pueden incluso generar nuevos problemas de gestión previamente no identificados.
- Su utilidad real se tambalea en ciertas ocasiones en las que una base datos *off-chain* es suficiente para resolver el dilema
- Se debe modificar la actual arquitectura de diseño hacia una híbrida (*on-chain/off-chain*) para cumplir con el RGDP, puesto que los bloques de transacciones no pueden almacenar datos personales.

En consecuencia, el problema desciende a nivel de almacenamiento de los datos personales, es decir, en las bases de datos fuera de los bloques de transacciones propuestos por la tecnología *blockchain*. Para enmarcar el estado de la situación y abordar el marco empírico se recuerdan los dos objetivos iniciales de la investigación de forma que nos permitan evolucionar en la misma:

- Objetivo 1: *Blockchain* es una tecnología que puede aportar una posible solución viable al problema de la falta de privacidad, confidencialidad y seguridad de los datos recolectados y usados en procesos de *Learning Analytics*.
- Objetivo 2: Es posible diseñar e implementar una solución para adecuar el nivel de confidencialidad de datos personales de alumnos impuesto por el Reglamento General de Protección de Datos durante el uso de *Learning Analytics* en Entornos Virtuales de Aprendizaje.

En este capítulo se expone el contexto empírico del problema abordado en esta tesis en toda su magnitud. La exposición transcurre en cinco partes distintas que explican los pasos realizados para la consecución del segundo objetivo. En primer lugar, se expone la metodología utilizada. En segundo lugar, se discute la necesidad de dar solución a las excepciones del RGPD en Moodle, terminando con una propuesta a desarrollar. En el

siguiente apartado se expone la ejecución de un proceso de experiencia de usuario mediante entrevistas semiestructuradas con el objeto de abordar el desarrollo funcional de la solución propuesta. El cuarto apartado expone el proceso de desarrollo de la solución. Finalmente, se muestran las líneas paralelas en las que se ha trabajado durante la tesis.

III.1. Metodología

El proceso de validación del segundo objetivo se desglosa en dos fases, una de epistemología cualitativa y la segunda enmarcada en una propuesta metodológica iterativa e incremental para desarrollar un prototipo evolutivo de una solución *software*. En la primera fase se realizan una serie de entrevistas con actores del contexto educativo, negocio, *Learning Analytics*, *blockchain* y legal para conocer posibles problemas relacionados con entornos virtuales de aprendizaje y RGPD. En esta primera fase se exponen:

- Las excepciones del RGPD descubiertas en Moodle.
- Un proceso de experiencia de usuario para recolectar percepciones sobre nuestra solución a una de las excepciones.

En la segunda fase se desarrolla un prototipo evolutivo que permite ejercer derechos legales específicos no disponibles, actualmente, en un entorno virtual de aprendizaje como es Moodle y detectados en la primera fase.

Se elige trabajar con Moodle como plataforma EVA donde crear los prototipos de los problemas detectados debido a que es la plataforma más usada en España. Según Hill (Hill, 2016), más del 65% de instalaciones de EVA en Europa son Moodle.

III.2. Excepciones del RGPD en Moodle

Los términos y condiciones de uso son un canal unidireccional, de la institución al estudiante, en el que este último puede aceptar o declinar el contrato. En caso de aceptar, el estudiante puede ejercer sus derechos. De entre los derechos otorgados por el RGPD, el derecho a la oposición de tratamiento plantea una cuestión esencial, ¿puede

un estudiante objetar en un curso, pero no en otro? Esto se resume en cuestiones de granularidad en el anonimato, de manera que un mismo estudiante pueda querer ser anónimo en un curso y en otro no. La granularidad en los EVA es una de las cuestiones que se debate en la investigación como camino a continuar.

La anterior cuestión abre una serie de conversaciones con el consultor especialista en cuestiones de protección y confidencialidad de datos personales asignado por La Salle Campus Barcelona. Se plantea la cuestión y se detecta que esta granularidad se considera una excepción fuera de lo común.

D. Amo: ...Solamente una última aclaración. En mi tesis indico que en Moodle un usuario puede aceptar distintos acuerdos legales y que estos acuerdos legales pueden estar asociados a un curso. Por tanto, podría ser que para cada curso el usuario tenga que aceptar un acuerdo legal con condiciones distintas. Entiende que el derecho a oposición le permitiría al usuario oponerse al tratamiento de un curso concreto ero permitirlo en otro, ¿o estoy equivocado?...

Consultor La Salle: ...No sería demasiado lógico que una institución tuviera condiciones de uso en espacios distintos que difirieran o se contradijeran... El único matiz que puede introducir el usuario es este derecho de oposición al tratamiento, que lo ejercen personas que tienen derecho a que su identidad sea preservada (por ejemplo, policías) o que hayan sufrido problemas de violencia de género... El derecho a la oposición al tratamiento, en su ejercicio, queda a criterio de cada persona, por tanto, podría exigir para un ámbito y no para otro ámbito..

(Consultor La Salle, comunicación personal, 20 de junio del 2019)

En la entrevista se evidencia que tal granularidad es posible y necesaria, pero que se consideran excepciones. No obstante, las excepciones deben ser solucionables. Para conocer el estado de la cuestión y cómo se aborda en Moodle se realizan dos pasos. El primero conocer qué funcionalidades dispone Moodle en cuanto al RGPD y el segundo contrastar el estado del tratamiento con un *partner* de Moodle.

III.2.1.1. *Cómo Moodle se adecua al RGPD*

Las soluciones aportadas por el propio Moodle en cuanto a la adaptación de su plataforma al RGPD son en formato de *plugins*. Por una parte, pone a disposición el *plugin Policias* (Téllez & Moodle HQ, 2018). Por otra parte, pone a disposición el *plugin Data Privacy* (Moodle HQ, Pataleta, & Monllaó, 2018). Ambos *plugins* resuelven algunos aspectos apuntados por el RGPD, sin embargo, tal y como se ha señalado, no permite que un usuario sea anónimo.

El *plugin Policias* proporciona:

- Un nuevo proceso de inicio de sesión de usuario.
- Definir múltiples políticas (sitio, privacidad, terceros).
- Realizar un seguimiento de los consentimientos del usuario.
- Gestionar las actualizaciones y versiones de las políticas.

El *plugin Data Privacy* proporciona:

- El flujo de trabajo para que los usuarios envíen solicitudes de acceso y borrado de datos.
- Opciones de procesado de las solicitudes de usuarios a los administradores del sitio y los delegados de protección de datos.
- Definir un periodo de retención de los datos almacenados en un sitio Moodle.

Los dos *plugins* ofrecidos por Moodle permiten las gestiones generales en cuestiones de protección de datos. No obstante, ofrecen una aproximación de todo o nada, donde los estudiantes solo tienen la opción de aceptar todas las condiciones si quieren entrar en la plataforma de aprendizaje. Las excepciones en Moodle no se pueden solucionar ya que no existe una funcionalidad que lo permita.

Tras analizar las soluciones aportadas por Moodle se realiza una entrevista con el *partner* de Moodle 3iPunt (“3iPunt,” 2019). En dicha sesión, se coincide en que Moodle no permite salvar las excepciones que plantea el RGPD (T. Llorach, comunicación personal, 3 de julio del 2019). En definitiva, es una característica no desarrollada en Moodle, los *partners* de Moodle son conscientes y se requiere una solución al respecto.

III.2.1.2. Propuesta de desarrollo

En consideración a la sección anterior, se propone una solución que resuelva el problema de la granularidad y excepciones en Moodle. Se selecciona desarrollar un *plugin* debido a las posibilidades de este formato en contraposición a otras, como sería la modificación de librerías del núcleo de Moodle o partes implicadas en el tratamiento de datos. Para el desarrollo del *plugin* se consideran las siguientes premisas:

- Modificar las bibliotecas del núcleo de Moodle crearía una versión alternativa de la plataforma, de la misma manera que modificar otras partes del código. Se quiere evitar modificar el código fuente de Moodle.
- No se desea crear una nueva plataforma, o un parche de actualización, sino un añadido que resuelva el problema de forma fácil, incluido para gestores de Moodle de un perfil técnico bajo.
- Un *plugin* permite:
 - Instalarse y desinstalarse a conveniencia por el administrador del sistema.
 - Instalarse en pasadas, presentes y futuras instalaciones de Moodle.
 - Una facilidad de mantenimiento superior del código a lo largo del tiempo.
 - Publicarse en el repositorio de Moodle y llegar a más instalaciones que sufran del problema.
 - Desarrollar un prototipo funcional en base a los dos *plugins* de privacidad de Moodle. Construir sobre unas funcionalidades ya testeadas es más rápido que desarrollar sin referencia.

Se añade que el formato *plugin* tiene mejor aceptación tanto para los administradores de Moodle como el mismo Moodle, que una modificación del núcleo. En definitiva, un *plugin* permite ser instalado y desinstalado a discreción de la institución educativa o del administrador.

El problema de la granularidad del RGPD en Moodle implica que un estudiante pueda ejercer el derecho de oposición al tratamiento en un curso, pero no en otro (ver III.2

Excepciones del RGPD en Moodle). La solución que se aporte debe abordar el tema de manera que:

- Se fácil de comprender por parte de los interesados (responsables de datos, delegado de protección de datos, institución educativa, administradores, profesores y estudiantes).
- Los interesados se sientan cómodos con la solución.
- La gestión y uso de la solución debe ser fácil de realizar por los interesados.

Se plantea otorgar a un estudiante una segunda identidad, puesto que permite que:

- Un estudiante tenga un alias que no lo identifique.
- Un estudiante actúe en el anonimato libremente y con normalidad.
- Profesores, responsables de los datos y delegados de protección de datos gestionen fácilmente las dobles identidades tanto dentro de Moodle como fuera.
- Se generen distintos niveles de anonimato, en los que incluso el profesor pudiera no saber la identidad real del estudiante que participa en su curso.

En este sentido, se ejecuta un proceso de experiencia de usuario para:

- Contrastar la idea del alias con agentes legales.
- Comprobar qué nivel de confianza confiere esta aproximación.
- Desarrollar el *plugin* a partir de la recolección de percepciones de distintos perfiles educativos.

III.3. Experiencia de usuario

Con el objeto de recoger las percepciones ante el uso de alias como protección del anonimato de usuarios en procesos de aprendizaje en línea, se opta nuevo por un enfoque cualitativo mediante el diseño y ejecución de una serie de entrevistas estructuradas.

Este proceso se enmarca dentro de las posibilidades y área de influencia del campo conocido como experiencia de usuario (Bevan, Carter, & Harker, 2015; Hassenzahl & Tractinsky, 2006). Se recogen aportaciones desde distintas perspectivas de usuario para

tomar una decisión de desarrollo en base a los resultados. Por consiguiente, se realizan diferentes entrevistas a distintos perfiles para asegurar un resultado afín a las observaciones y percepciones de usuarios en EVA. Este proceso consta de entrevistas de preguntas directas, cuestionarios en línea y finalmente de entrevistas tipo *Bipolar Laddering*, BLA (Pifarré & Tomico, 2007), una combinación mixta que ha demostrado previamente su validez científica en la evaluación de todo tipo de metodología y sistema aplicado a la educación (Fonseca, Pifarre, Redondo, Alitany, & Sanchez, 2013; Llorca, Zapata, Redondo, Alba, & Fonseca, 2018; Villagrasa, Fonseca, & Durán, 2014; Villagrasa, Fonseca, Redondo, & Duran, 2018). Esta primera aproximación va a servir para analizar la validez del uso de alias en relación con el RGPD.

III.3.1. Primera aproximación

A continuación, se presenta el guion de la entrevista que se conduce a expertos del ámbito legal relacionado con leyes de protección de datos personales.

III.3.1.1. Guion de entrevista

Se exponen en la Tabla 20 la estructura y en la Tabla 21 preguntas de la entrevista que se efectúa a los expertos sobre legalidad en cuestiones de protección de datos.

Tabla 20 Primera aproximación de encuesta a perfil legal

Criterio de entrevista	Descripción
Perfil de usuario experto	Abogado
Objetivo de la entrevista	Conocer la viabilidad del uso de alias para asegurar el anonimato de los estudiantes que así lo expresen en entornos virtuales de aprendizaje.
Tiempo aproximado	15-20 minutos

Tabla 21 Preguntas de la encuesta al perfil legal

Pregunta	Tipo de respuesta
-----------------	--------------------------

PA1. ¿Considera que el uso de alias en un entorno virtual de aprendizaje asegura un nivel adecuado de anonimidad?	Abierta
PA2. ¿Considera alguna otra solución más conveniente?	Abierta
PA3. ¿Considera que la tarea de gestionar la creación y asignación de alias bajo petición expresa del usuario debe ser asignada al Delegado de Protección de Datos?	Abierta
PA4. ¿Considera que el uso de algoritmos como <i>Deep Fake</i> para generar la información seudonimizada de los alias es un método válido?	Abierta

III.3.1.2. Resultados de la entrevista

La entrevista se realiza por teléfono. Su transcripción en la Tabla 22 resuelve las preguntas:

Tabla 22 Respuestas de la encuesta al perfil legal como primera aproximación

Pregunta	Perfil legal	Respuesta
PA1	1	Sí, muy bien es muy correcto
	2	Sí
	3	Si seudonomiza los datos personales ya es suficiente, sí.
PA2	1	No
	2	No se me ocurre ninguna
	3	El seudonomizado es suficiente, quizás otra alternativa complicaría el escenario
PA3	1	Sí, correcto
	2	Sí
	3	Sí, aunque sería posible delegar la gestión a otro responsable, pero en su defecto sí
PA4	1	Sí es correcto
	2	Sí, más que suficiente
	3	Se pueden utilizar distintos métodos. El <i>Deep Fake</i> sería válido

III.3.1.3. Conclusiones de la primera aproximación

Se considera que todo usuario de un EVA tiene el derecho a la anonimidad considerando lo expuesto en el Reglamento General de Protección de Datos (EP and the CEU, 2016), en materia de derechos de protección de datos de personas físicas y seudonimización, y en los puntos del reglamento 1, 6 y 28:

(1) La protección de las personas físicas en relación con el tratamiento de datos personales es un derecho fundamental. El artículo 8, apartado 1, de la Carta de los Derechos Fundamentales de la Unión Europea («la Carta») y el artículo 16, apartado 1, del Tratado de Funcionamiento de la Unión Europea (TFUE) establecen que toda persona tiene derecho a la protección de los datos de carácter personal que le conciernan.

(6) Los principios de la protección de datos deben aplicarse a toda la información relativa a una persona física identificada o identificable. Los datos personales seudonimizados, que cabría atribuir a una persona física mediante la utilización de información adicional, deben considerarse información sobre una persona física identificable.

(28) La aplicación de la seudonimización a los datos personales puede reducir los riesgos para los interesados afectados y ayudar a los responsables y a los encargados del tratamiento a cumplir sus obligaciones de protección de los datos.

Por consiguiente, el uso de alias para proteger el anonimato de un usuario en un entorno virtual de aprendizaje es una aproximación correcta para asegurar un adecuado grado de anonimato a los usuarios de EVA. Esto es debido a que:

- Un dato seudonimizado continúa siendo un dato de carácter personal que identifica a una persona.

- Por ende, su uso es correcto dentro de la legalidad vigente para asegurar un nivel adecuado de confidencialidad de los datos personales de los afectados.

Se destaca que en los entornos *Blended Learning* queda extraño conocer al estudiante en persona y encontrar una persona con nombre, apellidos y foto de alguien que no está en clase. No obstante, es importante considerar que un alias ayuda tanto a:

- Estudiantes que quieren mantener el anonimato por razones de casos particulares excepcionales, por ejemplo, violencia de género.
- Estudiantes que quieren ayudar aportando sus interacciones en el sistema, pero sin ser identificados.
- Profesores para crear nuevas oportunidades de aprendizaje donde en ciertos estudios se requieran actividades de *role-playing*.

En consecuencia, la solución a los alias tiene completo sentido tanto en aprendizajes en línea, presenciales y híbridos.

III.3.2. Encuestas a perfiles estudiantes

Las siguientes encuestas se orientan teniendo en cuenta las conclusiones y argumentos de III.3.1 Primera aproximación. En todas las encuestas se expone en detalle el contexto, el problema y la solución.

III.3.2.1. Guion de encuesta

Se expone en las Tablas Tabla 23 y

Tabla 24 la estructura y preguntas de la encuesta que se efectúa a los estudiantes.

Tabla 23 Criterios de encuesta a perfiles estudiante

Criterio de entrevista	Descripción
Perfil de usuario experto	Estudiante
Objetivo de la entrevista	Valorar la opinión de los estudiantes de usar alias en aprendizaje meramente en línea y mediado por entornos virtuales de aprendizaje para asegurar su anonimato
Justificación de la entrevista	Valorar la percepción del alumno ante el nuevo sistema

Método	Encuesta tipo formulario en línea
Tiempo aproximado	15-20 minutos

Tabla 24 Preguntas de la encuesta a estudiantes

Pregunta	Tipo de respuesta
PA1. ¿Te importaría presentarte con un alias ante los estudiantes de un curso en línea?	Sí / No
PA2. ¿Has usado alguna vez un alias para mantener el anonimato?	Sí / No
PA3. ¿Qué te parece usar un alias para salvaguardar tu anonimato en entornos virtuales de aprendizaje?	Likert 5 puntos

III.3.2.2. Resultados de las encuestas

La prueba se realiza de forma presencial o a distancia, siendo los estudiantes los que acceden al formulario en línea para responder a la encuesta. El tiempo promedio de las encuestas es de 15 minutos, que se utiliza para presentar la investigación, su contexto, la propia encuesta y responder a las preguntas. Cada estudiante responde de manera independiente y voluntaria, reflejándose las respuestas en las Figuras Figura 19, Figura 20 y Figura 21.



Figura 19 Respuesta a la pregunta PA1 de la encuesta a los estudiantes

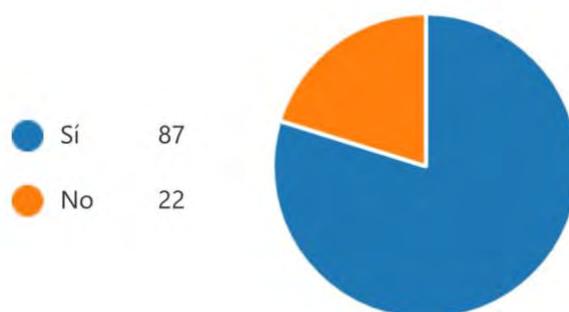


Figura 20 Respuestas a la pregunta PA2 de la encuesta a los estudiantes

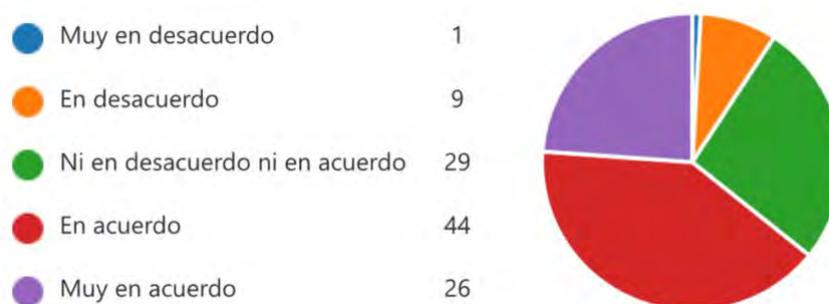


Figura 21 Respuestas a la pregunta PA3 de la encuesta a los estudiantes

III.3.3. Entrevistas a perfiles profesores

Las siguientes entrevistas se orientan teniendo en cuenta las conclusiones y argumentos de III.3.1 Primera aproximación. En todas las entrevistas se expone en detalle el contexto, el problema y la solución.

III.3.3.1. Guion de entrevista

Se expone en las Tablas Tabla 25 y Tabla 26 la estructura y preguntas de la entrevista que se efectúa a los profesores.

Tabla 25 Criterios de entrevista a perfiles profesor

Criterio de entrevista	Descripción
Perfil de usuario experto	Profesor
Objetivo de la entrevista	Valorar la percepción según la perspectiva del profesorado de usar alias en aprendizaje en línea y mediado por entornos virtuales de aprendizaje para asegurar su anonimato

Justificación de la entrevista	Valorar la percepción del profesorado ante el nuevo sistema
Método	BLA
Tiempo aproximado	15-20 minutos

Tabla 26 Preguntas de la entrevista a profesores

Pregunta
PP1. ¿El uso de alias complicaría la gestión de aula?
PP2. ¿Qué cree que aporta usar un alias como medida de anonimato?
PP3. ¿Qué otra solución cree se podría aplicar para resolver el problema del anonimato?
PP4. ¿Ha usado alguna vez un alias para mantener el anonimato?

III.3.3.2. Resultados de la entrevista

La entrevista se realiza de forma presencial. El tiempo promedio de las entrevistas es de 15 minutos. En cada una el entrevistado indica si tiene una percepción positiva o negativa, una valoración numérica del 0 al 10, una respuesta y comentarios de mejora si lo cree conveniente. Las Tablas Tabla 27, Tabla 28, Tabla 29 y Tabla 30 muestran las respuestas a cada una de las preguntas

Tabla 27 Respuesta a la entrevista a profesor PP1. ¿El uso de alias complicaría la gestión de aula?

Profesor	Percepción	Comentarios y mejoras
1	+ 10	Sin comentarios
2	+ 10	Sin comentarios
3	+ 8	Si fuera automática sería de 10, pero si la asignación de notas es manual entonces 8
4	+ 10	Si se presenta como alias no hay problema, el problema viene cuando no sabe que es alias y el alumno lo expresa
5	+ 8	Me da igual que se llamen como sea, son alumnos Mejoras: El que como profesor vea un alumno intermitente (concepto junta), puede no ayudar al alumno si no se sabe quién

		es. El profesor siempre debería saber si es un alias o no, sino no se puede evaluar con contexto (evolución, seguimiento, tutoría, etc.)
6	+ 8	Sería contraproducente tener una relación falsa Mejoras: El profesor debería saber quién es el estudiante
7	+ 10	Se mira el aprendizaje, que aquel usuario sea uno u otros es igual, se evaluará aquella persona con la interacción y resultado de las actividades Mejoras: A nivel de diseño, es diferente, ya que aceptar tener alias en el aula puede ser que no pueda hacer ciertas actividades
8	- 5	Sí me afectaría sobre todo en trabajos en grupo, si se tuviera que hacer una gestión multidisciplinar sería complicado porque no se estaría generando un grupo como se quisiera. Es negativo, pero no catastrófico. Por lo tanto, se debería diseñar y programar el curso de forma diferente Mejoras: Quizás restringiendo el perfil falso del estudiante, donde haya información real. Por lo tanto, si hay buenas restricciones del alias lo vería bien
9	+ 10	Yo creo estaría totalmente engañado en el caso de no saberlo, pero tampoco me implicaría nada
10	+ 10	No le veo ningún tipo de problema, personalmente me gusta el contacto directo, pero en un curso online lo veo indiferente. Tener un punto cercano está bien, pero el anonimato te permite prudencia, pero se pierde la proximidad
11	- 3	Depende lo que quieras hacer. No lo veo claro, me gusta que cuando empiezan clase se ponga foto. Doble: por una parte si tu carácter no es abierto te protege, pero por otra parte puede haber gente que se aproveche para hacer cosas indebidas
12	+ 10	Sí que complica... pensando por ejemplo en actas y notas
13	- 3	Sí ya que se pierde contexto del estudiante. Que el perfil sea idéntico para todas las asignaturas.

Tabla 28 Respuesta a la entrevista a profesor PP2. ¿Qué cree que aporta usar un alias como medida de anonimato?

Profesor	Percepción	Comentarios y mejoras
1	+ 9	Cree que la medida es adecuada, siempre que se le ofrezca un alias predefinido o a escoger entre opciones para evitar errores
2	+ 9	Lo ve muy positivo y no obligado
3	+ 10	Le parece excelente. Es lo más práctico
4	+ 10	Si alguien necesita conservar su intimidad, pues está en su derecho, imparto igual para unos que para otros. Los que tenemos delante son un número
5	+ 8	Por el concepto de no tener ideas preconcebidas, me fijo en el resultado y evolución independientemente de quién sea, dando objetividad dentro de una asignatura en concreto
6	+ 9	El alias debe ser serio, relacionado con la persona. Preferiblemente foto falsa, pero con seriedad (no Leo Messi), el género puede causar prejuicios, condicionar, etc.
7	+ 10	En general el uso del alias en el aula es negativo, pero para casos concretos justificados tiene sentido. En un contexto online, si todos fueran alias, entonces se perdería el <i>networking</i> y compartir experiencias enriquecedoras, todo sería como una mentira. A nivel genérico puede ser problemático y lo veo negativo, pero de forma puntual es justificable. Y debe poder ofrecer una protección. Esto conllevaría una diferenciación entre tipos de cursos
8	+ 6	Aporta confidencialidad de cara al estudiante y a los compañeros, pero quizás en el caso online puede crear desconfianza a los estudiantes, cuando hay trabajos colaborativos Mejoras: Quizás que los alumnos estén informados de que este usuario es un alias, tal vez esto ayudaría a confiar un poco más.
9	+ 10	Es una herramienta más positiva. Esto refleja un problema social, que se puede arreglar con alias.
10	+ 10	Es una medida súper simple a problemas complicados, no le veo ningún tipo de problema

		Mejoras: Garantizar un anonimato en un curso es un extra de marketing, teniendo la tranquilidad de que el estudiante se puede certificar
11	+ 1	Plantear las opciones negativas para su gestión
12	+ 10	Aporta soluciones para que el usuario pueda gestionarse de forma anónima
13	+ 8	Aporta independencia de resultados respecto la persona, se evalúan resultados

Tabla 29 Respuesta a la entrevista a profesor PP3. ¿Qué otra solución cree se podría aplicar para resolver el problema del anonimato?

Profesor	Percepción	Comentarios y mejoras
1	+ 10	El alias ya es suficiente y no veo ahora mismo otro sistema.
2	+ 10	Que todo el mundo aparezca como números y quien quiera que se haga público. Por defecto privado. El profesor puede saber los códigos y cómo se quiere que se puedan dirigir a ti. Que el propio estudiante se pueda crear el alias con sistema de validación (no sea nombre de famoso, etc.)
3	+ 10	El alias ya es suficiente
4	+ 10	Volvamos al número en lugar de nombre apellido y foto, puesto que puede coincidir con alguien ya existente Mejoras: Los números no deben poder repetir
5	+ 8	Estandarizaría el alias, que el alias no pueda cambiar y siempre sea el mismo Mejoras: El alias debería ser a-género, incluso la imagen debería ser un icono, para evitar sesgos/prejuicios en el profesorado e incluso alumno. Que el alias no sea personal
6	+ 10	O se identifica o se pone alias
7	+ 10	Por defecto el alias es negativo. Puede ser algo muy simpático para hacer algunas actividades, como hacer ciertas actividades de juegos de rol. El hecho de tener una herramienta de anonimato puede llevar a generar nuevas actividades formativas. Se puede usar como

		herramienta educativa, por ejemplo, un <i>rol play</i> (uso de Leo Messi por ejemplo o hacerse pasar por una persona)
8	+ 8	No haría falta hacer nada más que un alias Mejoras: Es un tema de restricciones
9	+ 10	El alias es suficiente para un curso totalmente online, es lo más práctico. Pero que alguien del centro sepa quién es quién, que la información real esté en algún sitio, también para tranquilidad para el estudiante en cuestiones de certificación Mejoras: A posteriori, 2 años después, estaría bien que supiera que había tenido tal estudiante. Para cuestiones de juicios/certificaciones debe quedar esta información registrada
10	+ 10	El alias ya es suficiente
11	+ 7	En ciertos momentos, anónimo en momentos puntuales, no para todo. Por ejemplo, solo para una actividad
12	+ 10	Como profesor, no podría por permisos, no se me ocurre cómo
13	+ 7	Crear un aula específica de casos anónimos, ya que puede haber la necesidad de los estudiantes de mantener un contacto real con otros estudiantes... incluso para establecer relaciones más personales.... Preguntar al inicio de la matrícula qué tipo de modelo de relación se quiere establecer

Tabla 30 Respuesta a la entrevista a profesor PP4. ¿Ha usado alguna vez un alias para mantener el anonimato?

Profesor	Percepción	Comentarios y mejoras
1	+ 10	Como profesor, necesitas tener máxima confianza en el administrador del sistema para validar que el usuario que estás validando es quien es. Para evitar fraudes. Como profesor certifico que este usuario supera el curso, pero no se puede decir que sea una persona
2	+ 10	No gusta mantenerse en el anonimato cuando alguien hace una crítica. Si no es por necesidad pienso que es mejor no utilizar el anonimato
3	+ 10	Sencillo y fácil de utilizar

4	+ 10	Parece bien el hecho de no tener que compartir datos personales identificativos. Es algo más rápido Mejoras: Crear alias modificable a partir de mis datos personales, si gusta me lo quedo sino lo cambio
5	+ 9	Lo veo como algo muy bueno. El concepto es que es más rápido y cómodo, incluso de niveles de seguridad (baja y alta)
6	+ 10	No, uso de redes sociales identificables. Quizá usaría la opción de oculto, y no sería en buenas intenciones. Aunque entiendo que puede haber circunstancias personales que se pueda utilizar
7	+ 5	Quizá por no tener la necesidad de utilizarlo no lo veo necesario, siempre soy transparente en la identidad. No lo percibo como negativo. Mi posición es neutral
8	+ 4	Nunca. Estoy fuera de redes. Es negativo para la parte de confianza en las redes sociales, si veo correcto que haya alias para casos específicos. El alias debe ser personalizable, sino no se adecuará a todos los programas, por tanto, tener un alias por curso tiene mucho sentido. Coordinado con secretaría, poner las restricciones que toque. El DPD pone mujer como alias y es lo que hay
9	+ 10	Sí, en cuentas de correo, por ejemplo, en los nombres de usuario Mejoras: Estaría bien poder escoger el alias e incluso personalizarlo
10	+ 10	Es imprescindible, depende de donde quieras ir es imprescindible. Hay gente que lo usa en acceso, porque desvirtúa la realidad. Con twitter no uso alias, depende del contexto creo que es necesario dar la cara, se pierde la credibilidad. Alguien que me critica dando la cara me puede dar mucha información de él, en cambio los que esconden lo único que hacen es incendiar con intereses concretos. Utilizaría aquellos servicios donde se identifica y abandonaría cualquiera que use <i>nicknames</i> . Pero encuentro muy interesante que haya alias / <i>nicknames</i> en ciertas ocasiones
11	+ 8	En entornos desconocidos para mantener mi independencia, en casos cercano no

12	+ 8	Depende del entorno
13	+10	Me adapto a la información que quiero pública en cada perfil

III.3.3.3. Resumen de las entrevistas

En la siguiente Tabla 31 se muestra la puntuación resumen de cada pregunta (-130 a +130) y la conclusión de los comentarios y mejoras.

Tabla 31 Resumen de la percepción de cada pregunta de las entrevistas a los profesores

Pregunta	Percepción	Resumen de comentarios
PP1	+ 83	En general no importa tener estudiantes con alias, puesto que se evaluará igual. No obstante, algunos profesores ven difícil realizar actividades afines al grupo clase cuando el profesor no sabe que existe un alias entre los estudiantes. La percepción de gestionar un aula con usuarios alias es notablemente positiva
PP2	+ 110	En general ven el alias como una solución válida y muy sencilla para resolver problemas sociales. Algún profesor indica que si todo fueran alias se perdería la autenticidad en ciertas actividades. Se repite la idea de generar alias de información discreta. La percepción de usar un alias en el aula es muy positiva
PP3	+ 120	Ningún entrevistado ofrece una solución alternativa y en general se apunta a que el alias es más que suficiente para solucionar problemas de anonimato. Algunos profesores apuntan a que el alias sea inmutable, que sean números y que el profesor sepa quién es quién. La percepción de que no exista una solución mejor que un alias es muy positivo
PP4	+ 66	En general se considera que siempre las personas deberían identificarse, pero que el alias es un recurso útil para resolver problemas muy concretos. La percepción de usar un alias en general es menos positiva

III.3.4. Entrevistas a perfiles administradores técnicos de Moodle

Las siguientes entrevistas se orientan teniendo en cuenta las conclusiones y argumentos de III.3.1 Primera aproximación. En todas las entrevistas se expone en detalle el contexto, el problema y la solución.

III.3.4.1. Guion de entrevista

Se expone en las Tablas Tabla 32 y Tabla 33 la estructura y las preguntas de la entrevista que se realiza a los administradores técnicos de Moodle.

Tabla 32 Criterios de entrevista a perfiles administradores técnicos de Moodle

Criterio de entrevista	Descripción
Perfil de usuario experto	Administrador técnico Moodle
Objetivo de la entrevista	Valorar las inquietudes técnicas en relación con el uso de alias y su posible aplicación como <i>plugin</i> en los entornos virtuales de aprendizaje gestionados
Justificación de la entrevista	Valorar la percepción del administrador técnico ante el nuevo sistema
Método	BLA
Tiempo aproximado	15-20 minutos

Tabla 33 Preguntas de la entrevista a los administradores técnicos de Moodle

Pregunta
PAT1. ¿Cree que un <i>plugin</i> es una solución técnica compatible con vuestro entorno virtual de aprendizaje?
PAT2. ¿Cree que un <i>plugin</i> de estas características es la solución más adecuada?
PAT3. ¿Qué otra solución complementaria cree que se podría desarrollar?
PAT4. ¿Ha usado alguna vez un alias para mantener el anonimato?

III.3.4.2. Resultados de la entrevista

La prueba se realiza de forma presencial. El tiempo promedio de las entrevistas es de 15 minutos. En cada una el entrevistado indica si tiene una percepción positiva o negativa, una valoración numérica del 0 al 10, una respuesta y comentarios de mejora si lo cree conveniente. Se muestra en las Tablas Tabla 34, Tabla 35, Tabla 36 y Tabla 37 las respuestas a las preguntas.

Tabla 34 Respuesta a la entrevista a administrador PAT1. ¿Cree que un plugin es una solución técnica compatible con vuestro entorno virtual de aprendizaje?

Administrador	Percepción	Comentarios y mejoras
1	- 3	Sí, aunque dudo de su uso. Prefiero que si es norma, que esté de base, prefiero no instalar parches
2	+ 10	Sí
3	- 2	No, todo <i>plugin</i> necesita de instalación y peligra en situaciones de actualizaciones. Todo integrado en los paquetes o sistemas utilizados
4	+ 7	Sí, supongo. Utilizamos google drive, existen complementos, supongo que es viable... peligro las actualizaciones
5	+ 9	Sí, ¿en caso de migración a nuevas versiones se contempla actualización y adaptación?

Tabla 35 Respuesta a la entrevista a administrador PAT2. ¿Cree que un plugin de estas características es la solución más adecuada?

Administrador	Percepción	Comentarios y mejoras
1	+ 10	En la plataforma sí. Si es aceptado institucionalmente sí... y si es un derecho que esté incluido en una de las versiones
2	+ 10	Complicado. Quizás con permisos se podría solucionar y no sería tan complejo
3	+ 8	Quizás. En caso de necesidad es tan sencillo como instalarlo y ejecutarlo, debe ser una opción disponible en función de necesidad

4	- 3	No, un <i>plugin</i> tiene la necesidad de adaptarse a navegador, sistema operativo, etc., mejor si estuviera integrado de forma genérica... no lo veo claro
5	+ 8	No me lo he planteado, supongo que sí por la opción de instalarlo o no..., pero no me queda claro el proceso para el DPD

Tabla 36 Respuesta a la entrevista a administrador PAT3. ¿Qué otra solución complementaria cree que se podría desarrollar?

Administrador	Percepción	Comentarios y mejoras
1	- 4	Se utilizan otras herramientas. Se tendría que adaptar a otros entornos... como Facebook, Whatsapp, etc. Necesidad de que el alias fuera multiplataforma
2	+ 10	Con permisos, se puede excluir que los alumnos vean toda la información
3	+ 10	Subgrupos en función de características de los usuarios y/o perfiles. Hay que atender a la diversidad y a las necesidades concretas de cada estudiante
4	+ 8	Que en el proceso de matrícula del estudiante se genere directamente la protección necesaria, con un potencial <i>flag</i> solo visible en caso de autorización del estudiante para su visualización por el profesor. Importante tener diversas soluciones posibles en función de los sistemas que se utilicen
5	+ 7	La aplicación de control multisistema. Necesidad de controlar multisistemas y plataformas

Tabla 37 Respuesta a la entrevista a administrador PAT4. ¿Ha usado alguna vez un alias para mantener el anonimato?

Administrador	Percepción	Comentarios y mejoras
1	- 5	Sí. En Tripadvisor, lo usé por vergüenza, aunque no me gusta
2	+ 8	No. Depende del entorno

3	+ 10	Sí. Me gusta separar mis perfiles profesionales de los de ocio y dentro de estos según sistema para que no se vinculen actividades
4	+ 8	No. Solo utilizo sistemas que me interesan, en los que no me hace falta anonimizarme
5	+ 10	Sí. Es bueno para independizar usos y contenidos

III.3.4.1. Resumen de las entrevistas

En la Tabla 38 se muestra la puntuación resumen de cada pregunta (-50 a +50) y la conclusión de los comentarios.

Tabla 38 Resumen de la percepción de cada pregunta de las entrevistas a administrador técnico de Moodle

Pregunta	Percepción	Resumen de comentarios
PAT1	+ 21	En general la percepción de usar un <i>plugin</i> es una solución compatible. No obstante, existe una clara preocupación en las posibles actualizaciones del sistema y consecuente actualización del <i>plugin</i> . Hay apreciaciones en las que se desea una implementación del <i>plugin</i> dentro del sistema de manera que se ofrezca por defecto
PAT2	+ 33	La percepción del uso de un <i>plugin</i> como solución más adecuada es positiva, debido a su sencilla instalación y desinstalación. Se proponen alternativas como la posibilidad de que una configuración de permisos pueda dar los mismos resultados
PAT3	+ 31	Se proponen soluciones complementarias al <i>plugin</i> , por ejemplo, la necesidad de desarrollar en otras plataformas y convertir el alias en multiplataforma, que se atiendan las necesidades según las características del estudiante o que en el proceso de matriculación pueda realizarse la configuración pertinente. Se sigue con la idea de utilizar permisos como alternativa factible al uso de alias
PAT4	+ 31	Hay una contraposición en el uso del alias, aunque su uso en dependencia del entorno se percibe como positivo. Algunas respuestas reflejan que no gusta utilizar alias, por lo que

solamente se usan sistemas en los que no es necesario usar alias.
Otras respuestas reflejan un uso de alias a conveniencia en
dependencia del entorno

III.3.5. Entrevistas a perfiles Delegados de Protección De Datos

Las siguientes entrevistas se orientan teniendo en cuenta las conclusiones y argumentos de III.3.1 Primera aproximación. En todas las entrevistas se expone en detalle el contexto, el problema y la solución.

III.3.5.1. Guion de entrevista

Se expone en las Tablas Tabla 39 y Tabla 40 la estructura y preguntas de la entrevista que se efectúa a los delegados de protección de datos (DPD).

Tabla 39 Criterios de entrevista a perfiles Delegados de Protección de Datos

Criterio de entrevista	Descripción
Perfil de usuario experto	Delegado de Protección de Datos
Objetivo de la entrevista	Valorar la comodidad ante el proceso de generación y asignación de alias a los estudiantes en entornos virtuales de aprendizaje que así lo requieran de forma expresa. Teniendo en cuenta que la comunicación con el usuario sigue canales estándares ya establecidos
Justificación de la entrevista	Valorar la percepción del delegado de protección de datos
Método	BLA
Tiempo aproximado	15-20 minutos

Tabla 40 Preguntas de la entrevista a delegado de protección de datos

Pregunta
PD1. ¿Cree que es el perfil adecuado para gestionar las peticiones de anonimato?
PD2. ¿Cree que es el perfil adecuado para gestionar los alias de usuario?

PD3. ¿Qué otra solución cree que se podría aplicar para resolver el problema del anonimato?

PD4. ¿Ha usado alguna vez un alias para mantener el anonimato?

III.3.5.2. Resultados de la entrevista

La prueba se realiza de forma presencial. El tiempo promedio de las entrevistas es de 15 minutos. En cada una el entrevistado indica si tiene una percepción positiva o negativa, una valoración numérica del 0 al 10, una respuesta y comentarios de mejora si lo cree conveniente. Se muestra en las Tablas Tabla 41, Tabla 42, Tabla 43 y Tabla 44 las respuestas a las preguntas.

Tabla 41 Respuesta a la entrevista a administrador PD1. ¿Crees que eres el perfil adecuado para gestionar las peticiones de anonimato?

DPD	Percepción	Comentarios y mejoras
1	- 4	No lo sé, no lo tengo claro, no considero necesario, me cuesta comprender los casos en que podría decir que sí o que no
2	+ 9	Sí. No obstante, depende el numero de solicitudes se puede generar problemática Mejora: tener claro opciones de limitar
3	+ 10	Sí. Soy la encargada de mi escuela de conocer la normativa legal y posibles soportes tutoriales al estudiante, donde entiendo que encaja el sistema explicado
4	+ 10	Sí, puesto que soy el que recibe todas las peticiones concernientes a datos de los estudiantes

Tabla 42 Respuesta a la entrevista a administrador PD2. ¿Crees que eres el perfil adecuado para gestionar los alias de usuario?

DPD	Percepción	Comentarios y mejoras
1	- 4	No, se debieran derivar al servicio técnico
2	- 3	No, es sobrecargar el trabajo de un perfil Mejora: buscar soporte en los servicios técnicos del curso
3	+ 9	Sí. Al ser la encargada de la oficina de datos y gestionar el enlace con los profesores, puedo actuar de forma independiente

4	+ 9	Mi departamento es el encargado de gestionarlo, pero no de crear los alias
---	-----	--

Tabla 43 Respuesta a la entrevista a administrador PD3. ¿Qué otra solución cree que se podría aplicar para resolver el problema del anonimato?

Profesor	Percepción	Comentarios y mejoras
1	+ 10	Siempre que lo soliciten generar un alias o algo similar de forma inicial y permanente en el tiempo
2	+ 10	Aplicando restricciones a contenidos y usuarios
3	+ 9	Nada a comentar
4	+ 10	Es suficiente

Tabla 44 Respuesta a la entrevista a administrador PD4. ¿Ha usado alguna vez un alias para mantener el anonimato?

DPD	Percepción	Comentarios y mejoras
1	+ 7	No, supongo que puede ser útil en entornos generalistas y para evitar <i>spam</i>
2	+ 7	No, útil para usos masivos, pero no para usos limitados donde es mejor crear interacciones personales
3	+ 10	Sí, independizo usos según plataforma
4	+ 10	No, nunca uso alias, aunque considero que hay casos en los que a alguien le puede ser conveniente usarlo

III.3.5.3. Resumen de las entrevistas

En la Tabla 45 se muestra la puntuación resumen de cada pregunta (-40 a +40) y la conclusión de los comentarios.

Tabla 45 Resumen de la percepción de cada pregunta de las entrevistas a los Delegado de Protección de Datos

Pregunta	Percepción	Resumen de comentarios
PAT1	+ 25	En general la percepción de gestionar las peticiones de anonimato es positiva. No obstante, hay quien duda de que se puedan gestionar todas las peticiones
PAT2	+ 11	La percepción de gestionar los alias de los usuarios es positiva, sin embargo, se cree conveniente tener disponible el soporte técnico

		de la administración para balancear cualquier posible sobrecarga. Se apunta que la gestión es posible, pero no la creación de alias
PAT3	+ 39	En general se percibe el uso de alias como una solución positiva en la que las alternativas derivan a una posible configuración de los permisos y accesos a contenidos
PAT4	+ 34	En general no se considera necesario usar alias para mantener el anonimato, no obstante, se considera positivo su uso en casos en los que puede ser necesario, como evitar spam o servicios masivos

III.3.6. Conclusiones de los resultados

Desde la perspectiva de la legalidad, es válido usar un alias para proteger el anonimato de un usuario en un EVA. A pesar de la validez de los alias, es necesario definir la percepción relativa a la confianza que esta aproximación genera en los distintos roles de usuarios de un EVA. De lo contrario, esta solución podría caer en el mismo contexto de desconfianza que *Learning Analytics*. Precisamente se quiere encontrar una solución que permita el anonimato a cualquier estudiante, pero que también sea un procedimiento de confianza y no genere recelos o dudas en el proceso.

En las entrevistas a los estudiantes se destaca que un gran porcentaje se siente cómodo al utilizar un alias (84%) y muchos ya lo han utilizado (76%). Asimismo, un alto porcentaje de estudiantes se sentiría cómodo al usar un alias en un curso en línea (69%).

En las entrevistas a los profesores no existe ninguna percepción negativa. El uso del alias se percibe como positivo y genera confianza para resolver problemas sociales. En general no les importa tener estudiantes con alias, lo ven como una solución válida y muy sencilla, sin necesidad de encontrar otra solución. Algunas apreciaciones indican que podrían verse dificultades o limitaciones en algún tipo de actividad de aula. No obstante, estos escollos son salvables en el momento que el profesor sabe que existen alias entre los estudiantes.

En las entrevistas a los administradores se destaca que el uso de alias se percibe como positivo en aquellos entornos que lo requieran, pero siendo algunas respuestas contrarias al uso de alias como norma. El uso de alias a conveniencia en dependencia

del entorno se percibe como conveniente. En cuanto al formato de solución, ya sea en distribuible como *plugin* o como funcionalidad del producto, hay una tendencia hacia la segunda. Esta posición radica en el hecho lógico de es preferible una funcionalidad propuesta desde el producto puesto que tiene soporte directo del proveedor y evolucionará en consonancia a las actualizaciones del producto. A pesar de esta tendencia, se considera positivo utilizar la solución en formato *plugin* debido a su sencillez de instalación o desinstalación. Como alternativa se propone poder configurar el sistema para solucionar el problema, no obstante, en Moodle no existe la posibilidad de crear una configuración adecuada tal y como desean los administradores.

En las entrevistas a los delegados de protección de datos se destaca un bajo uso de alias para mantener el anonimato, pero se considera positivo su uso en casos en los que puede ser necesario, como evitar *spam*. En general se percibe el uso de alias como una solución positiva, no obstante, existe la incertidumbre de poder gestionar un gran volumen de peticiones y se hace referencia al soporte técnico como posible ayudante. Se extrae que es aceptable gestionar las peticiones y la asignación de alias por el delegado de protección de datos, incluso junto a su equipo o departamento, pero no así la creación de los usuarios alias.

Tras el análisis de las entrevistas conducidas a los cinco tipos de perfiles de usuarios, se concluye que asignar un alias a un usuario de un EVA genera suficiente confianza como para desarrollar un prototipo funcional en formato *plugin*.

III.4. Desarrollo de una solución en formato *plugin*

Retomando el hilo de la investigación, es necesario recordar la definición del problema de la granularidad como aquel donde un estudiante expresa querer ser anónimo en un curso y en otro no. En un curso de Moodle todos los estudiantes ven quién está matriculado y también su perfil público con datos personales. Por consiguiente, y mediante un *plugin*, debe habilitarse una opción que permita ejecutar el derecho de anonimato e impedir que los usuarios de un curso vean aquellos que no deseen ser vistos.

El desarrollo del *plugin* se basa en el precepto de proteger el anonimato de los usuarios a la vez que se cumpla con el RGPD. Por este motivo se asigna el nombre “*Protected users*”, para dar a entender que existen usuarios protegidos en el Moodle que requieren de especial atención.

El *plugin* desarrollado y presentado en la investigación es funcional. No obstante, en cuestiones de diseño se toma la aproximación “recorrido cognitivo” (Granollers, Perdrix, & Lorés, 2004). De esta manera se realiza el diseño primigenio de las pantallas antes de pasar a una fase de validación de la usabilidad. La validación de la usabilidad es un proceso largo que se traslada a trabajos futuros como continuación de la tesis doctoral. El diseño del *plugin* sufre dos evoluciones antes de ser funcional y estar preparado para una primera liberación. Se procede a exponer su evolución y estado final.

III.4.1.1. Primer diseño

En un primer estadio se propone asignar un rol especial a los estudiantes de manera que puedan configurar aquella información que no se quiera mostrar públicamente. A continuación, se muestran los diseños de las distintas pantallas. No obstante, este procedimiento se desecha por estar en contradicción con las premisas de desarrollo. Para que esta versión sea funcional se debe modificar el código fuente de Moodle y en las premisas de desarrollo se explicita que no sea así.

En la Figura 22 se muestra el flujo de funcionamiento del *plugin* en su primer diseño.

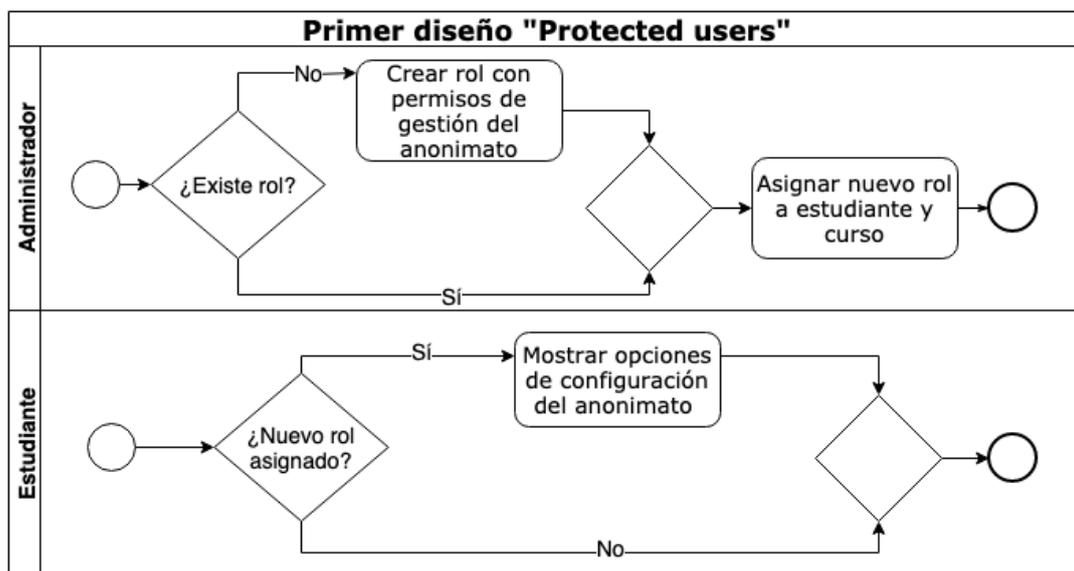


Figura 22 Diagrama de flujo de funcionamiento del primer diseño del plugin "Protected users". Fuente: Elaboración propia.

- El administrador puede crear un rol de estudiante con los nuevos permisos de gestión (ver Figura 23 y Figura 24).
- El administrador asigna el rol al usuario y al curso en el que quiere ocultar datos personales (ver Figura 25).
- El usuario, y desde la cabecera del curso, gestiona las opciones de anonimato de sus datos personales (ver Figura 26 y Figura 27).

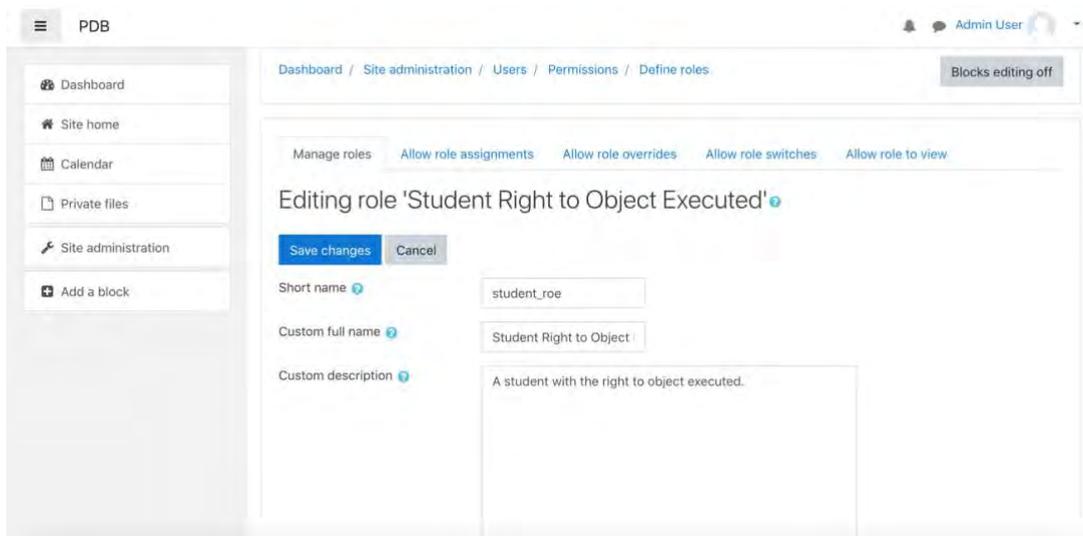


Figura 23 Edición del rol de usuario que permite a un estudiante gestionar el anonimato de sus datos privados. Fuente: Elaboración propia.

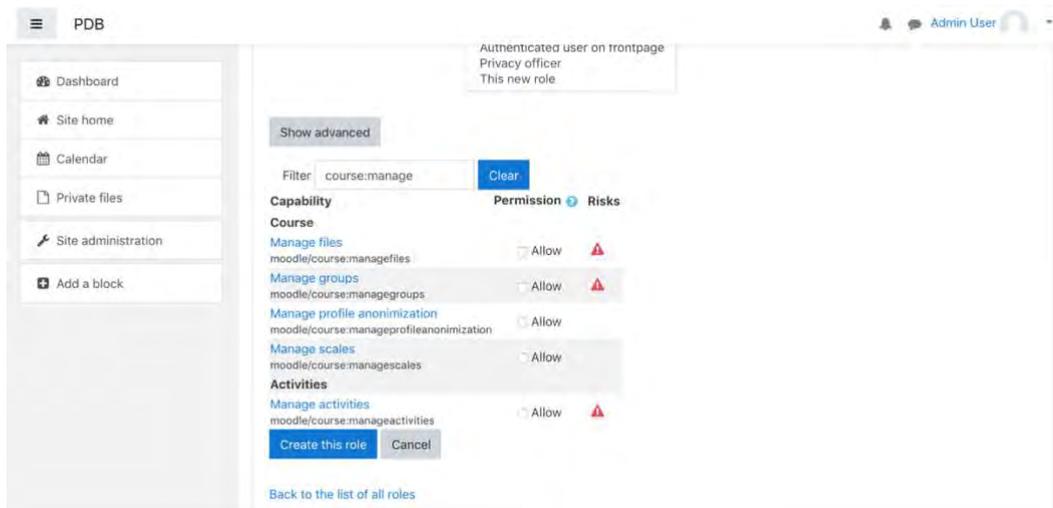


Figura 24 Asignación de permisos al rol de usuario que permite a un estudiante gestionar el anonimato de sus datos privados. Fuente: Elaboración propia.

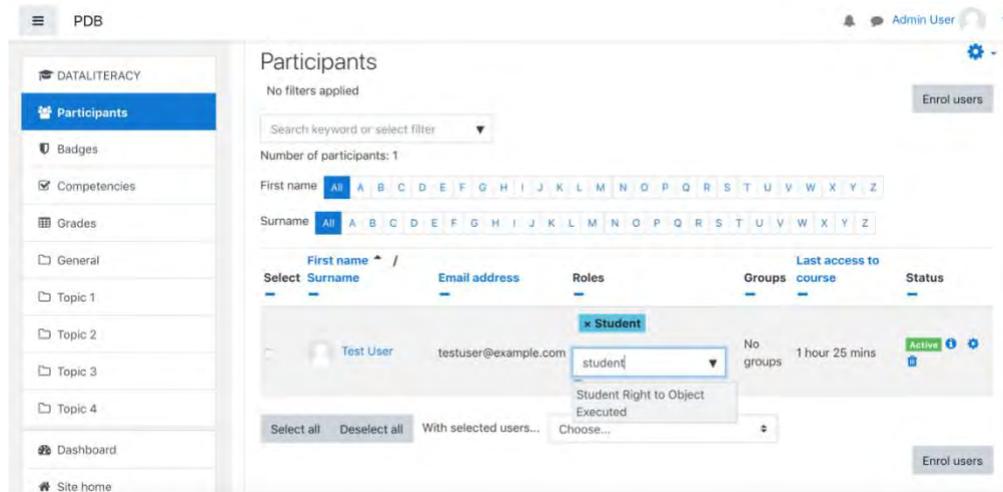


Figura 25 Asignación del rol con permisos de gestión de anonimato al estudiante de un curso en concreto. Fuente: Elaboración propia.

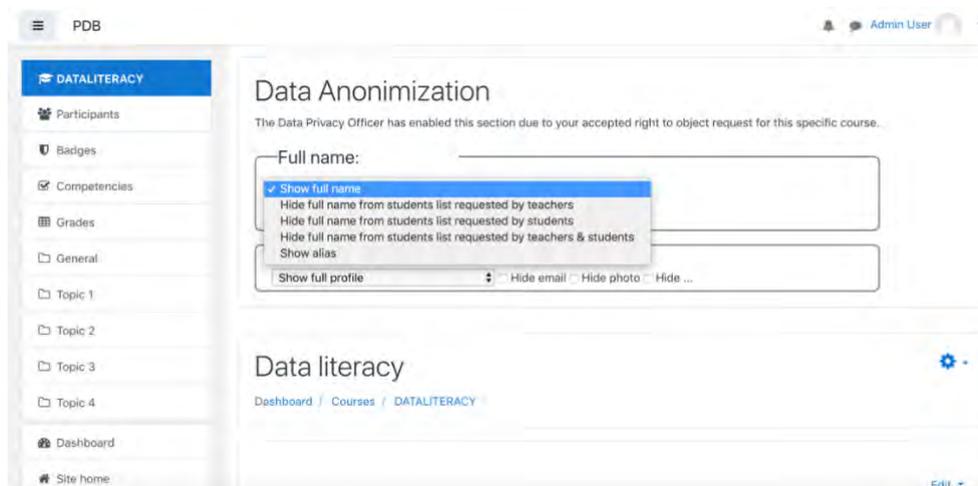


Figura 26 El usuario puede gestionar qué datos ocultar desde la cabecera del curso (I). Fuente: Elaboración propia.

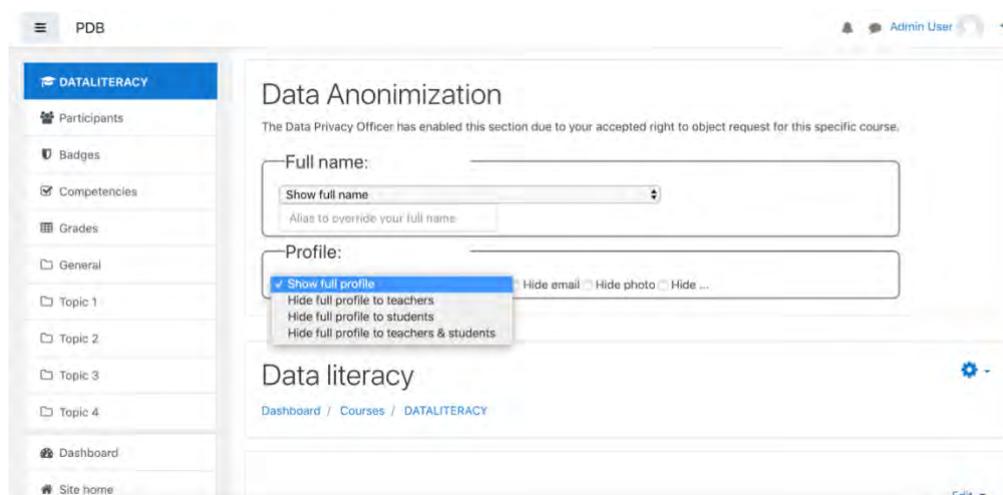


Figura 27 El usuario puede gestionar qué datos ocultar desde la cabecera del curso (II). Fuente: Elaboración propia.

III.4.1.2. Primera evolución

En un segundo estadio se elimina la excesiva granularidad en la configuración del anonimato y se vuelve al problema inicial. Se hace foco en asignar un alias a un usuario de manera que pueda conectarse con este a cualquier curso. Los motivos técnicos están expuestos en el punto anterior.

En este estadio se toma una serie de decisiones que consisten en:

- Eliminar los permisos y roles asignables a los estudiantes.
- Crear un usuario alias para los usuarios protegidos.
- Mover la responsabilidad de la gestión de alias al delegado de protección de datos.
- Que un usuario pueda acceder a sus alias desde la página principal de la plataforma una vez esté conectado.

Las decisiones expuestas permiten al usuario, considerado usuario protegido, expresar voluntariamente al delegado de protección de datos el requerimiento de anonimato en el curso que lo desee. Es el delegado de protección de datos que crea un alias y lo asigna al usuario protegido. El usuario protegido puede entonces conectarse automáticamente a cualquiera de sus alias desde su perfil de usuario Moodle. En la Figura 28 se muestra el flujo de funcionamiento del *plugin* en esta primera evolución:

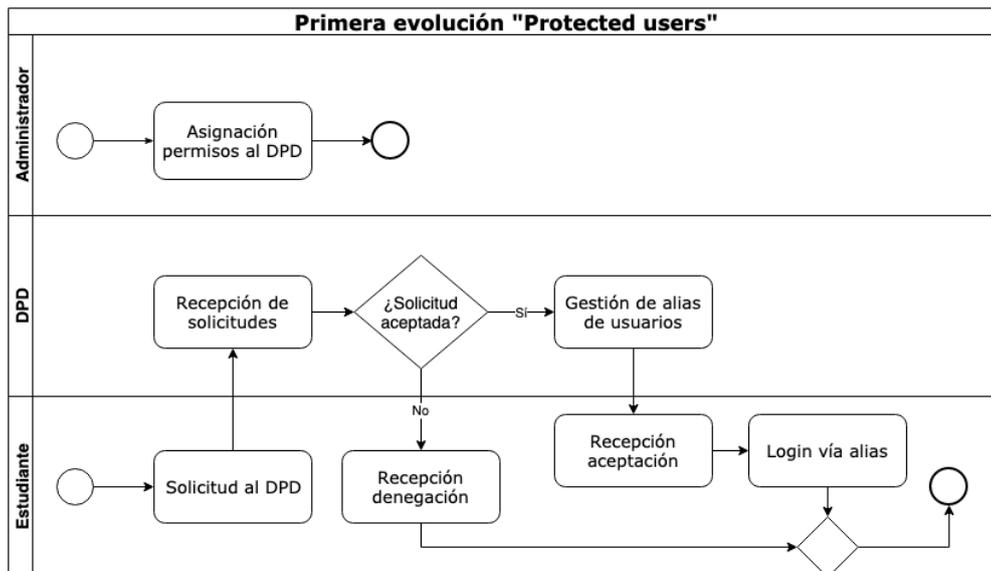


Figura 28 Diagrama de flujo de funcionamiento de la primera evolución del plugin "Protected users". Fuente: Elaboración propia.

- El administrador asigna permisos de gestión de alias al delegado de protección de datos.
- El usuario por proteger utiliza el formulario de contacto con el delegado de protección de datos para enviar la petición de anonimato (ver Figura 29).
- El delegado de protección de datos gestiona las peticiones enviadas, estudia el caso y las valida o rechaza (ver Figura 30 y Figura 31).
- El delegado de protección de datos asigna distintos alias al usuario por proteger (ver Figura 32).
- El usuario protegido accede a sus alias desde la página principal (ver Figura 33).

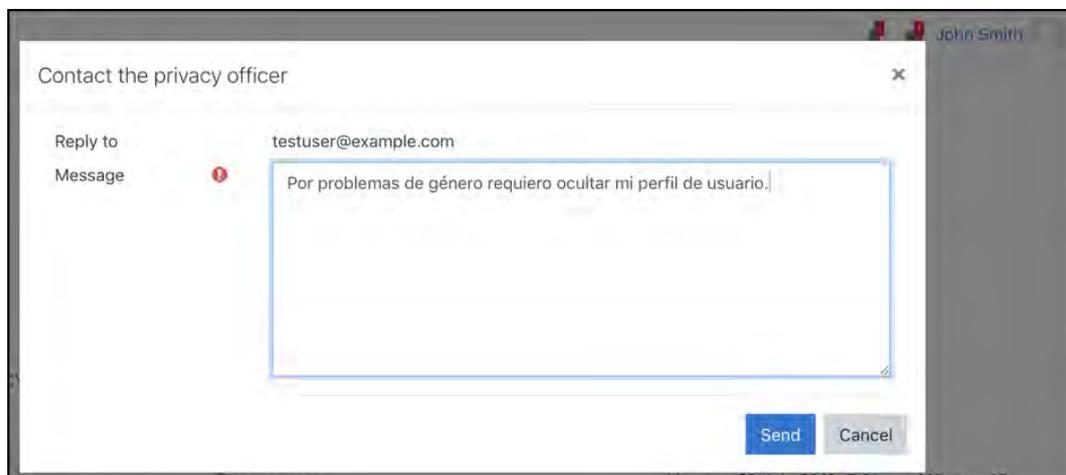


Figura 29 Formulario de contacto con el delegado de protección de datos. Fuente: Elaboración propia.

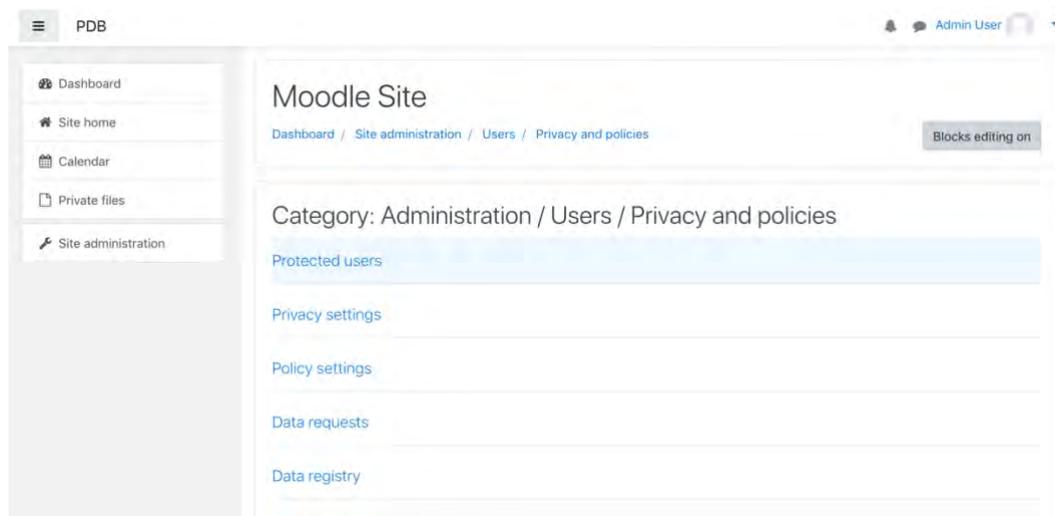


Figura 30 Acceso a protected users como delegado de protección de datos. Fuente: Elaboración propia.

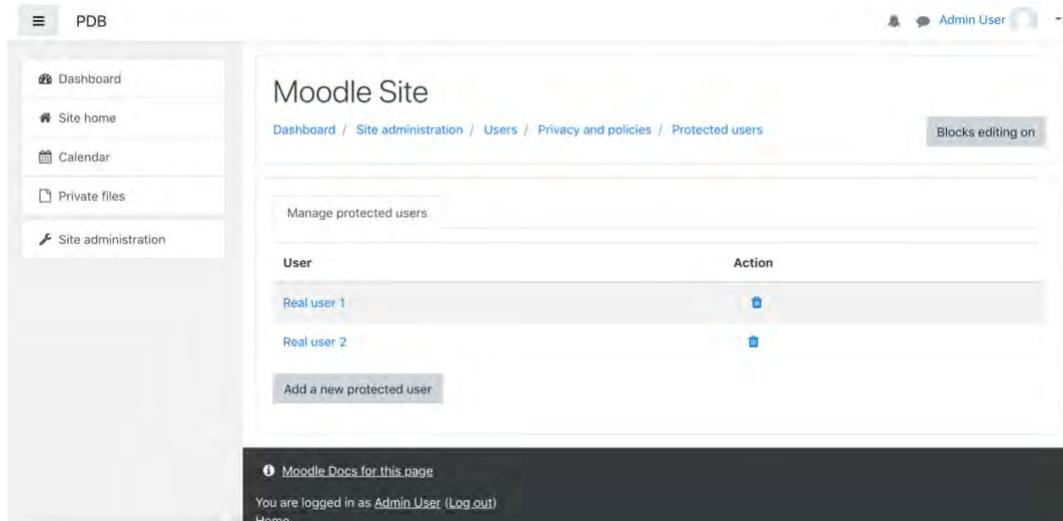


Figura 31 Gestión de los alias de usuarios protegidos versión 1. Fuente: Elaboración propia.

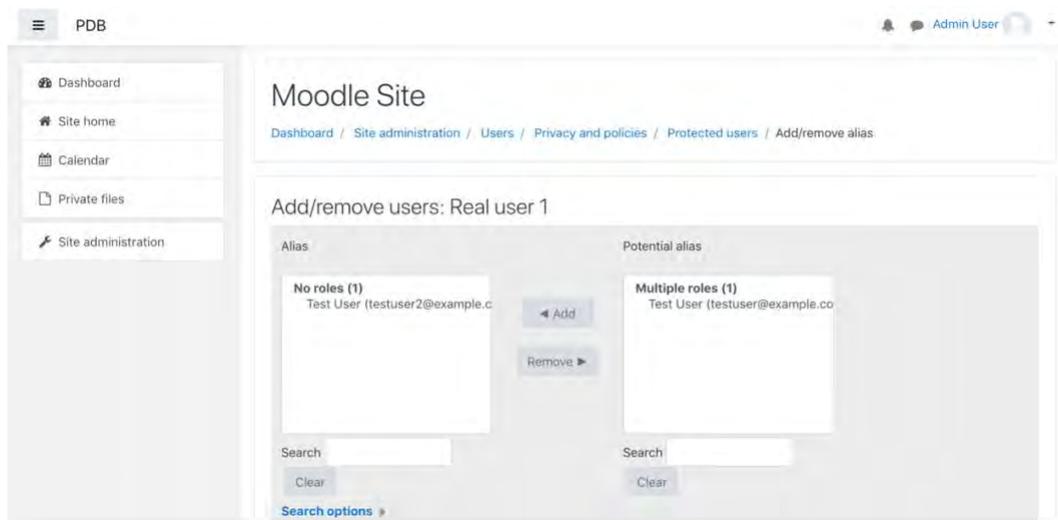


Figura 32 Asignación de alias versión 1. Fuente: Elaboración propia.

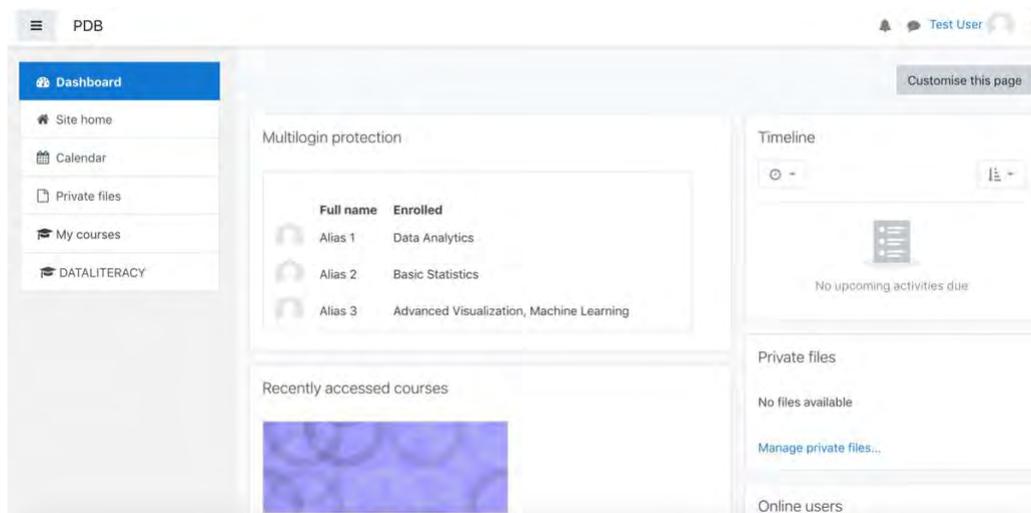


Figura 33 Acceso a alias desde página del usuario protegido. Fuente: Elaboración propia.

III.4.1.3. Segunda evolución

En un tercer estadio se realizan distintos cambios para mejorar la interfaz y el proceso de gestión, asignación y acceso a alias. Se elimina la intermediación del administrador, se facilita la gestión de los alias al delegado de protección de datos y se facilita el acceso a los alias del usuario protegido desde su perfil de usuario. Se procura mostrar la información lo más sencilla, minimalista y clara posible. La evolución de la interfaz se termina en esta segunda evolución.

La Figura 34 muestra el flujo de funcionamiento del *plugin* en esta segunda evolución.

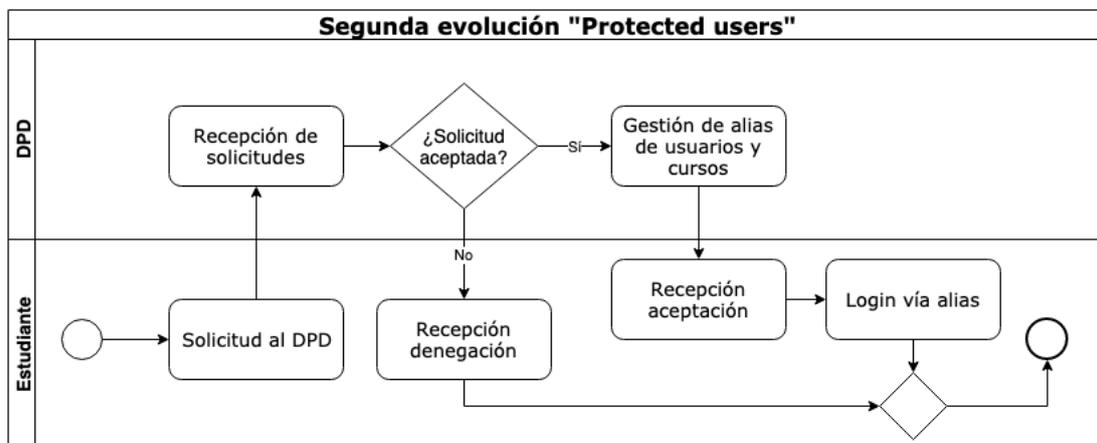


Figura 34 Diagrama de flujo de funcionamiento de la segunda evolución del plugin "Protected users". Fuente: Elaboración propia.

- El usuario por proteger utiliza el formulario de contacto con el delegado de protección de datos para enviar la petición de anonimato (ver Figura 29).
- El delegado de protección de datos gestiona las peticiones enviadas, estudia el caso y las valida o rechaza (ver Figura 35).
- El delegado de protección de datos asigna los alias al usuario en una visualización en formato informe (ver Figura 37). Se puede enlazar al alias directamente en un curso o múltiples cursos.
- El usuario protegido accede a sus alias desde la página principal (ver Figura 38 y Figura 39).

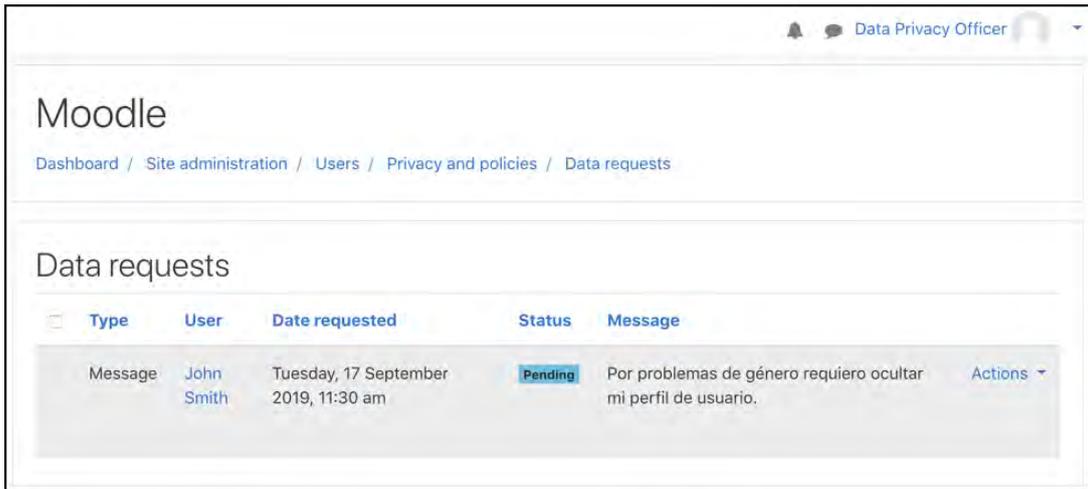


Figura 35 Gestión de peticiones por parte del delegado de protección de datos. Fuente: Elaboración propia.

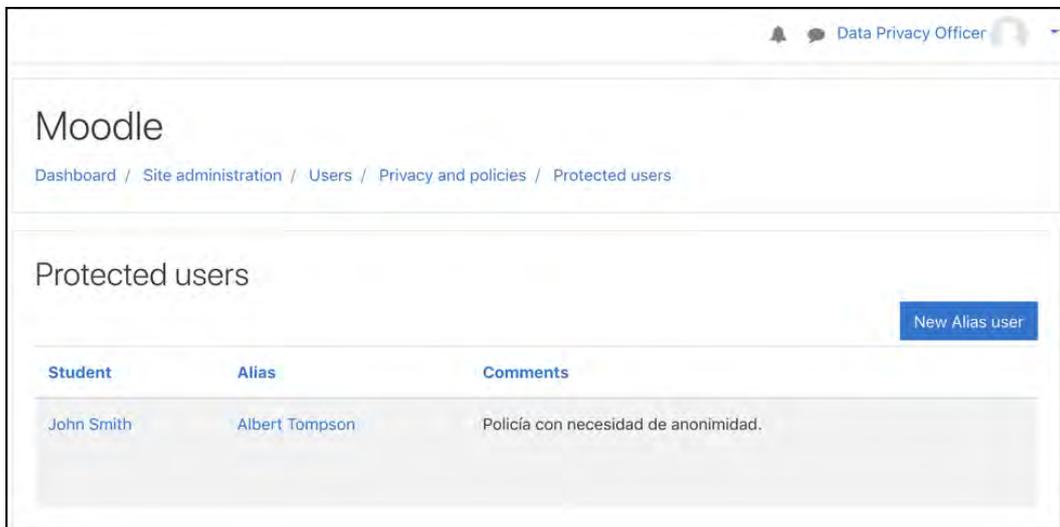


Figura 36 Gestión de alias de usuario versión 2. Fuente: Elaboración propia.

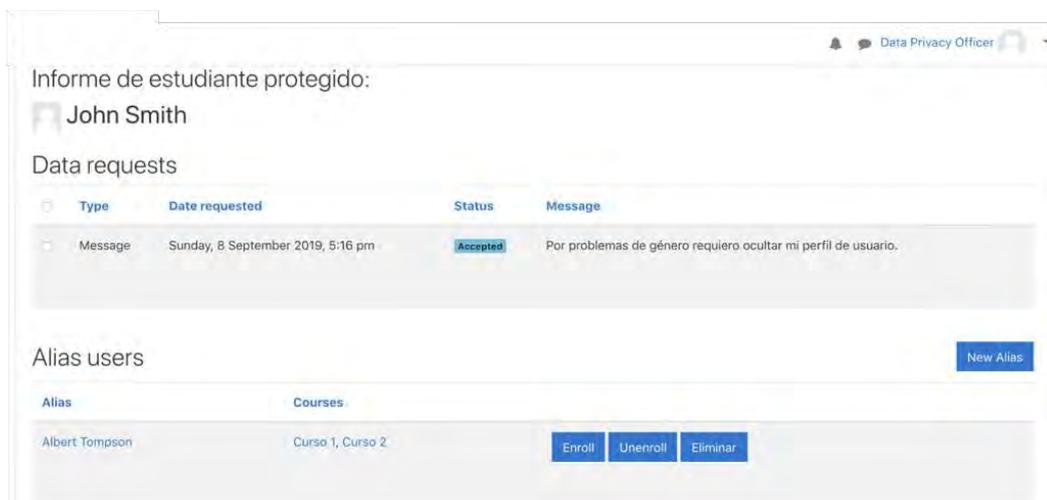


Figura 37 Informe administración del alias de un usuario protegido. Fuente: Elaboración propia.

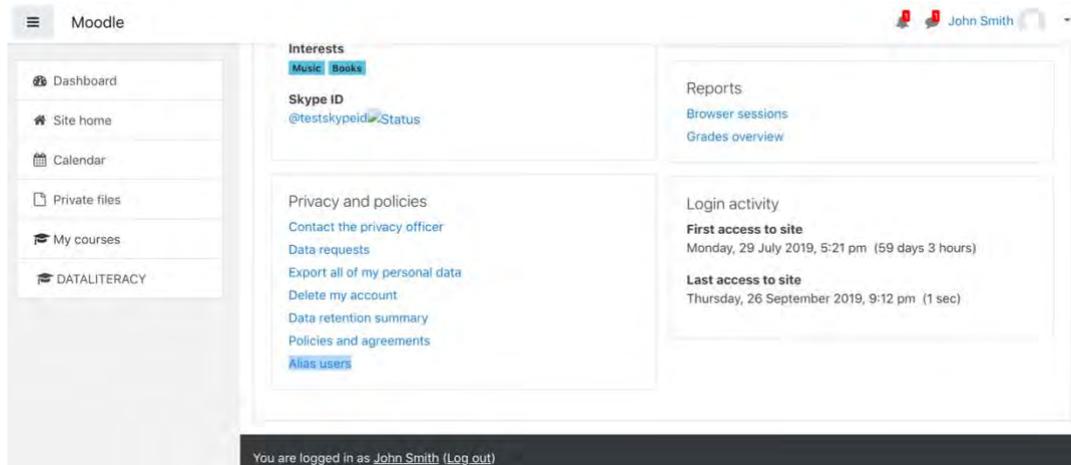


Figura 38 Acceso a alias del usuario protegido desde el perfil de usuario. Fuente: Elaboración propia.



Figura 39 Lista de multialias a los que puede acceder el usuario protegido. Fuente: Elaboración propia.

El siguiente paso consistiría en validar la usabilidad del *plugin*. El objetivo de la investigación es desarrollar un prototipo funcional. Se consigue tras el desarrollo del *plugin* expuesto y la ejecución de un recorrido cognitivo para resolver cualquier incidencia en el flujo de funcionamiento. Se deja el proceso de validación de la usabilidad como líneas futuras de trabajo.

No obstante, la usabilidad e interés del *plugin* tienen una primera prueba funcional positiva al haber sido seleccionados para su presentación, y potencial implantación en Moodle, en la Moodelmoot Global 2019 (Alier & Amo, 2019). Esta conferencia internacional es de gran importancia puesto que es el encuentro de la comunidad Moodle en la que se congregan agentes de peso de distintas partes del mundo. Por consiguiente, se confirma que el trabajo realizado es de gran interés y de gran calado.

Se siguen pasos de registro como modelo de uso en vistas de su potencial integración con Moodle.

III.5. Líneas trabajadas en paralelo

Solucionar problemas a nivel de bases de datos con datos personales y en entornos virtuales de aprendizaje implica dos aproximaciones:

- Por un lado, proteger y asegurar un nivel de confidencialidad de los datos personales adecuado al RGPD.
- Por otro lado, asegurar un nivel de seguridad de los datos almacenados adecuado al RGPD.

En el Marco empírico se muestra cómo abordar las excepciones relacionadas con el derecho a la oposición dentro del problema principal y no solucionada en EVA, como Moodle. Se demuestra que en esta cuestión es posible asegurar un nivel de confidencialidad de los datos personales mediante un alias (ver III.4 Desarrollo de una solución en formato *plugin*). La solución consiste en un desarrollo que permite a un usuario conectarse como alias en el curso en el que haya pedido anonimato de forma expresa. Se añade que es una solución interesante para distintos perfiles de usuario.

Asimismo, se trabaja paralelamente en otras soluciones, usando Moodle como plataforma de soluciones prototipo, para abordar distintas cuestiones dentro del problema principal. Estas cuestiones se relacionan con la seguridad en la tabla de usuarios, en los registros de interacciones y en los datos almacenados.

III.5.1. Seguridad en la tabla de usuarios

La tabla de usuarios de Moodle es el centro de operaciones de la plataforma. Todo acceso se relaciona con esta tabla y todo dato personal sale de la misma. Asegurar la tabla de usuarios es esencial para proteger la identidad de los alumnos y asegurar un nivel adecuado de seguridad en casos de accesos indebidos.

En un entorno donde la institución y el administrador de la plataforma se conocen y existe un alto elevado de confianza es necesario tomar o incrementar medidas de

seguridad y privacidad ante posibles ataques externos, puesto que en entornos de confianza es de donde proceden. Las posibilidades de encriptación en una base de datos MySQL de Moodle son varias. Se decide utilizar funciones *hash* SHA-2, puesto que es el estándar más nuevo, potente y disponible de las familias de funciones *hash* en MySQL (Preneel, 2010; Sobti & Geetha, 2012). A la vez, y aunque las funciones *hash* son muy adecuadas para seudonimizar datos personales, se realiza una capa adicional de encriptación utilizando otro estándar disponible en MySQL como es el AES (*Advanced Encryption Standard*) (Liberatori, Otero, Bonaadero, & Castiñeira, 2007; Popa, Redfield, Zeldovich, & Balakrishnan, 2012). La función *hash* y la función de encriptación se aplican en la tabla de usuarios, de manera que la seudonimización se refuerza; quedando así la tabla de usuarios blindada ante accesos indebidos. Se realiza la encriptación mediante un proceso de dos pasos:

1. *Hash* con la función SHA2 de 512 bits de longitud.
2. Una encriptación posterior con el algoritmo AES de 128 bits.

Con una pequeña modificación en el módulo de sesiones de Moodle, se crea para cada usuario, una tabla de los usuarios relacionados con los cursos a los que está inscrito. De esta manera un *hacker* que pueda hacerse con las llaves de acceso de un usuario solamente tendrá acceso a los usuarios de los compañeros de los cursos a los que está inscrito. Se añade que los datos de esta nueva tabla están encriptados con un proceso SHA2-512+AES-128.

Paralelamente, se crea un juego de vistas que revierte a tiempo real la encriptación de la información. De esta manera los datos están seguros en disco y en caso de acceso indebido se limitaría el impacto al afectar solo a una porción de usuarios de toda la base de datos, relacionados con el usuario explotado. La Figura 40 muestra el esquema de funcionamiento de la solución propuesta para proteger la tabla de usuarios de Moodle.

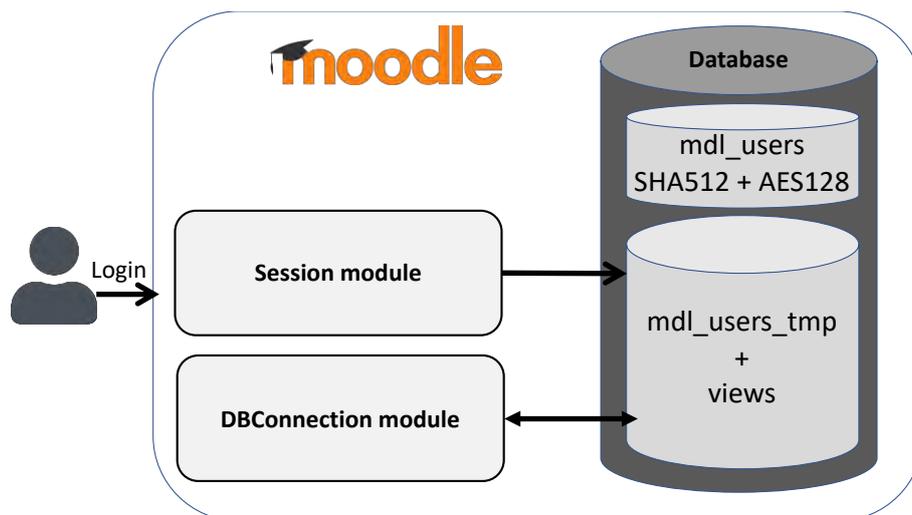


Figura 40 Diagrama de protección de la tabla de usuarios de Moodle. Elaboración: propia.

Se descarta seguir trabajando en esta solución debido a la demostración de sus resultados positivos, pero a los inconvenientes que presenta si el acceso administrador al sistema gestor de base de datos se ve comprometido.

III.5.2. Seguridad en los registros de interacciones

En un EVA se recolectan todas las interacciones de los estudiantes y se almacenan en registros. Esta fase se encuadra en la recolección de datos en procesos de *Learning Analytics*. Posteriormente, estos datos se tratan, analizan y visualizan para ayudar a los profesores a tomar mejores decisiones educativas basadas en datos. La seguridad en este proceso es clave para proteger las identidades de los estudiantes.

En el marco del trabajo (Amo, Fonseca, Alier, García-Peñalvo, Casañ, et al., 2019) se desarrolla un *plugin* de Moodle que permite almacenar los registros de interacciones en una base de datos fuera del dominio del administrador. En las aplicaciones de la tecnología *blockchain*, el eslabón más débil es la capacidad del usuario de no perder la clave privada. En un EVA, el eslabón más débil es el administrador del sistema y su capacidad de acceso a todo el conjunto de datos. En este sentido, el *plugin Moodle Personal Data Brocker Log Store* (Amo & Alier, 2019) permite externalizar las interacciones de los estudiantes a otro servidor fuera del alcance del administrador de la plataforma. Al mismo tiempo, y como el *plugin* es de código abierto, pueden añadirse tantas capas de seguridad como se crean convenientes, por ejemplo, encriptar los datos

desde origen para evitar su transferencia en abierto. La Figura 41 ejemplifica el funcionamiento del *plugin*.

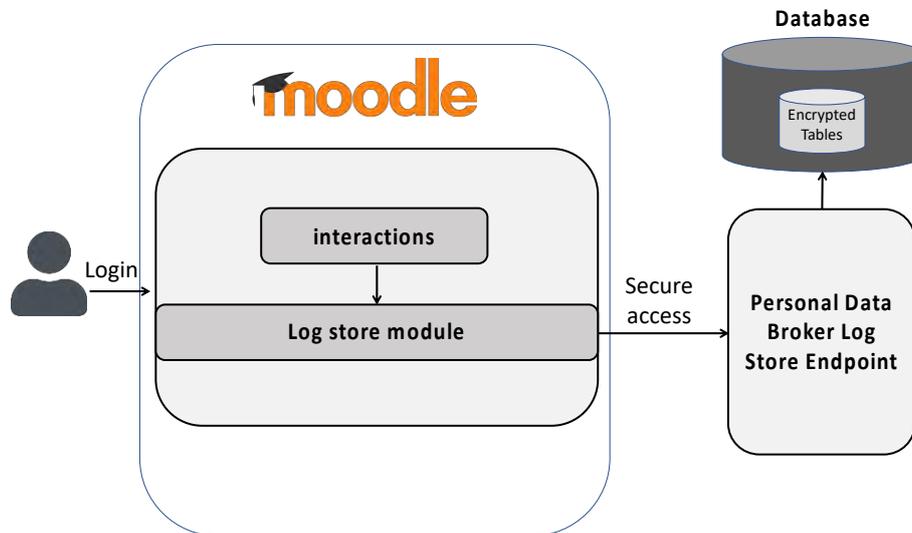


Figura 41 Diagrama del plugin Personal Data Broker Log Store. Elaboración: propia.

Se ha descartado seguir con esta aproximación debido a los inconvenientes que presenta en entornos de poca confianza. El administrador del sistema puede modificar el *plugin* instalado y derivar los datos a una base de datos paralela sin que nadie se percate. En un contexto donde no hay confianza se requiere de un nivel de seguridad superior, donde una implementación de la tecnología *blockchain* no puede hacer frente.

III.5.3. Seguridad en los datos almacenados

En el marco del congreso TEEM 2019 se presenta una nueva solución más acorde a contextos en los que la confianza entre administradores e instituciones educativas es muy baja (Amo, Alier, García-Peñalvo, Fonseca, & Casañ, 2019). Es una solución extrema que surge de intentar mitigar las cuestiones de granularidad en el cumplimiento del RGPD. Se pretende añadir una capa adicional de autorización a nivel de controlador del Sistema Gestor de Base de Datos (ver Figura 42). De esta manera solo aquel usuario autorizado por el EVA a acceder a datos arbitrarios será también autorizado por el Sistema Gestor de Base de Datos (SGBD).

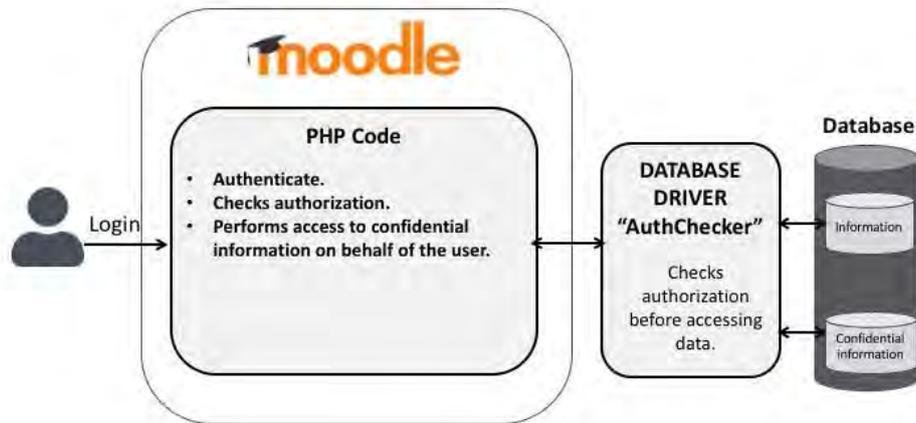


Figura 42 Diagrama de AuthChecker, controlador intermedio entre Moodle y almacenamiento de datos. Fuente: Elaboración propia.

El proceso de desarrollo implica considerar una matriz de acceso a datos que defina inicialmente quién tiene acceso, a qué datos tiene acceso y en qué contextos está autorizado, e incluso cuánto tiempo tiene permitido acceder a los datos. Esta matriz de acceso a datos es lo que ofrece la tecnología *blockchain*, pero en formato de base de datos relacional.

La solución presentada obliga al SGBD a consultar la matriz de acceso a datos. De este modo, cada vez que se quiere acceder a datos de la base de datos, se valida si el usuario está autenticado en el EVA y su sesión está activa. En caso contrario se deniega el acceso. Los datos están encriptados a nivel de disco duro y el SGBD devuelve los datos en abierto si su autenticación es positiva. Es una solución en aquellos entornos en los que no hay confianza y se requiere una seguridad máxima. Asimismo, es necesario un socio tecnológico que pueda hacer frente al desarrollo de la solución, terminándose su continuidad durante la investigación y dejándose como trabajo a futuro.

La Figura 43 muestra una comparativa de las tres soluciones propuestas.

positivas	negativas	
facilidad	semiautomático sobrecarga	seguridad tablas Moodle
seguridad	adaptación desarrollo complejo	PDB log store
máxima seguridad	máxima complejidad desarrollo muy complejo	AuthChecker

Figura 43 Cuadro comparativo de las líneas trabajadas. Elaboración: propia.

III.5.4. Interoperabilidad con *blockchain*

En el transcurso de la tesis se asesora a un estudiante de máster en el desarrollo de su Trabajo Final de Máster “El uso de *Blockchain* para resguardar analíticas de aprendizaje” (Rodríguez, 2019). En este trabajo se demuestra que es posible almacenar los permisos de acceso a los datos guardados en la base de datos de un EVA, pero no almacenar los datos en el *blockchain* de manera efectiva. *Blockchain* puede funcionar como puerta de acceso a datos y por lo tanto facilitar la interoperabilidad entre herramientas (Bdiwi et al., 2017).

Otra de las promesas de los interesados en crear soluciones con la tecnología *blockchain* es la privacidad interna de la identidad del usuario. Por ende, supuestamente se preserva el anonimato de los usuarios. ¿Puede entonces una implementación de la tecnología *blockchain* preservar la identidad de los estudiantes y asegurar un nivel de confidencialidad y seguridad de datos según lo establecido en el RGPD? Una respuesta afirmativa daría solución al problema de la investigación. Pero la respuesta es un “quizás” para algunos autores (Millard, 2018) y un “es complicado” para otros (Bacon et al., 2018), tanto por cuestiones técnicas como por cuestiones legales. Asimismo, los autores coinciden en que solo una solución implementada con tecnología *blockchain*, de acceso privado y con permisos podría llegar de manera teórica a cumplir con la legalidad.

Sin embargo, un desarrollo basado en tecnología *blockchain*, de acceso privado y con permisos puede no ser necesario en cuanto al diseño y requisitos del proyecto tecnológico. Se pueden conseguir los mismos resultados usando un SGBD en lugar de una implementación de la tecnología *blockchain*. En el momento que debe usarse una implementación de la tecnología *blockchain*, de acceso privado y con permisos para asegurar el cumplimiento del RGPD, se pone en entredicho la necesidad de un sistema *blockchain* (Narayanan, 2015). Por consiguiente, en un sentido puramente tecnológico, no tiene sentido usar *blockchain* cuando se puede usar, por ejemplo, una base de datos

relacional (Wust & Gervais, 2018). En tales casos, *blockchain* añade complejidad innecesaria.

En relación con el debate *blockchain*/SGDB, se realiza en una entrevista con el CEO (*Chief Executive Officer*) de la empresa ORKEI. Se llega a la conclusión de que el uso de la tecnología *blockchain* permite evidenciar fraude y por lo tanto funciona como un registro supuestamente inmutable. No obstante, son características que se pueden conseguir con un SGDB. La tecnología *blockchain* aporta un protocolo de consenso, pero este protocolo también puede aplicarse a una base de datos relacional:

D. Amo: Por tanto, estamos hablando de unas trazas de *logs* y hacer el seguimiento fuera de la base de datos. Evidentemente, una interfaz es necesaria para ver los datos tanto dentro de *blockchain* como en la base de datos, por ello, todo pueden ver las transacciones ya se use *blockchain* o base de datos. Como se puede solucionar la no trazabilidad de *logs* en un sistema gestor de base de datos, aunque esto suponga que tiene un coste elevado, entonces ¿con una base de datos sería suficiente? Entonces la decisión de usar *blockchain* o no viene dada por los costes y no por la tecnología, ya que entiendo que ambas soluciones son buenas.

E. Céspedes: En esencia tienes razón, pero piensa que conceptualmente en un *blockchain* no guardas el estado final sino todo un conjunto de transacciones que te dan un estado como resultado de ejecutarlas todas. Este es en esencia un sistema de *Event Sourcing*. El añadido que da *blockchain* sobre un sistema de *Event Sourcing* (que también existe en formato de base de datos: <https://eventstore.org/>) es que el *blockchain* encadena una transacción (o bloque de transacciones) con la anterior. Precisamente por este encadenamiento es por lo que necesitamos algoritmos de consenso.

D. Amo: Muy interesante... Por tanto, para concluir, un grupo de entidades pueden decidir utilizar 1) un eventstore.org basado en base de datos 2) *blockchain* si quieren una validación de transacciones por consenso. ¿Sería esto?

E. Céspedes: sí, generalizando sí

(E. Céspedes, comunicación personal, 4 de septiembre del 2019)

Las conclusiones de las entrevistas mantenidas con el estudiante del Trabajo Final de Máster (Rodríguez, 2019) son, incluso, más contundentes y en relación al RGDP:

D. Amo: ...llego a las siguientes dos conclusiones que espero puedas comentar. 1) Un *blockchain* privado, con permiso y que cumpla con el RGDP rompe con la idea inicial de *blockchain*. Por lo tanto, no debe llamarse *blockchain*, ya que a) no es necesario PoW, se puede usar PoS pero se rompe con el principio de confianza, b) se puede añadir, modificar y eliminar información, c) se rompe la descentralización. Este tipo de *blockchain* es más parecido a una lista ordenada por tiempo gestionada por un ente central. 2) Considerando el supuesto anterior, una base de datos sería suficiente para resolver la problemática que necesite un *blockchain* privado y con permisos donde la información se puede añadir, modificar y eliminar.

E. Rodríguez: Si bien es cierto que al subir datos personales a una *blockchain*, ya sea de forma pública o privada, viola los principios de la RGDP al ya no permitir a los usuarios remover su información; esta situación se evita completamente con guardar la información *off-chain*, y, en efecto, una *blockchain* modificada, que permita eliminar información ya procesada, iría en contra del propósito del *blockchain* porque la cadena ya se no podría recrear. Además, tengo dos consideraciones. 1) Las llaves públicas no deben de ser consideradas como información personal, ya que esa llave, a diferencia del correo electrónico no fue hecha directamente por el usuario. Dicho esto, la problemática que plantea el estudio queda anulado a partir de esta premisa; y 2) Si se llegase a tener un *blockchain* en el que se pueda eliminar o editar información de los bloques pasados, sería solo una base de datos distribuida que requiere un procesamiento extra para operar y perdería la confianza que *blockchain* tiene en este momento.

(E. Rodríguez, comunicación personal, 30 de septiembre del 2019)

De tales entrevistas se concluye que una implementación de la tecnología *blockchain* opera a un nivel por encima de un sistema gestor de bases de datos, donde actúa como una base de metadatos complementaria, por ejemplo, para resolver de otra manera aspectos de interoperabilidad o de certificaciones.

IV. Conclusiones y trabajos futuros

Una vez finalizados los procesos descritos en el marco teórico y en el empírico, dirigidos a conseguir los objetivos de la investigación mediante la resolución de los objetivos específicos identificados, a continuación, se da paso a discutir los resultados obtenidos. Las conclusiones discurren a modo de disertación a lo largo de los distintos capítulos y apartados, con un discurso de unión de los puntos más relevantes de la investigación. En la disertación se darán las referencias a la resolución de las preguntas y objetivos específicos relacionados con cada objetivo general, pudiéndose acceder al segmento concreto del manuscrito consultando las tablas disponibles en el apéndice 0 Localizaciones de las respuestas a preguntas y consecución de objetivos de los Apéndices. Por último, se muestran las líneas de trabajo futuras, las publicaciones y transferencias de conocimiento realizadas a lo largo de la investigación.

IV.1. Conclusiones

A lo largo del manuscrito se ha enmarcado que Internet está originalmente desprovisto de privacidad (ver sección II.2.1.1 Internet). Es un aspecto muy importante que considerar, puesto que la Red de Redes se creó pensando en un ecosistema abierto y transparente, el cual se considera como parte de la génesis del problema presentado en la investigación (ver sección I.1 Acotación del objeto de estudio). En la misma línea, las tecnologías web nacieron copiando la idea de apertura y comunidad, llegando a una situación distópica de captura de datos continuada de los cibernautas del siglo XXI (ver sección II.2.1.2 Web: HTTP y HTML), incluidos los estudiantes de cualquier edad.

Como se presenta en la sección II.2 Internet insegura y *clickstream*, el contexto Internet-Web-plataformas-*apps* es sensible debido a la fácil exposición de los datos privados de los usuarios. Ante este contexto desprovisto de privacidad, tanto instituciones educativas como empresas tecnológicas cuyo modelo de negocio se basa en *Big Data*, usan y construyen plataformas web de aprendizaje con metodologías de captura de datos como el *clickstream* (ver sección II.2.2 *Clickstream*). Por una parte, extraen comportamientos que pueden utilizarse para mejorar sus servicios y beneficiar a los

estudiantes. Por otra parte, pueden tomar decisiones estratégicas para ser más competitivos e incluso hacer de *data brokers* realizando acciones en contra de los intereses de sus clientes. En conclusión, los datos personales de los estudiantes, incluidos los de menores, se recopilan, se almacenan en abierto y fuera de control en distintos lugares y se tratan con fines supuestamente educativos, y lo que socava así la privacidad y seguridad de los datos recolectados. Se genera una inseguridad que levanta incógnitas como ¿quién tiene estos datos?, ¿quién los almacena? ¿quién y cómo los trata?, ¿en qué momento?, ¿para qué fin?, ¿las entidades responsables cumplen la legalidad? o ¿con quién se comparten estos datos? (ver sección II.4.4 Miedos y recelos: una cuestión delicada).

El análisis de datos de los estudiantes en EVA se realiza mediante procesos de *Learning Analytics* (ver origen en secciones II.3 MOOC y II.4 *Educational Data Mining*). En estos procesos analíticos la exposición de la privacidad aumenta y se vuelven más sensibles cuando participan menores de edad, cuya voluntad de participación queda supeditada a las herramientas usadas en el aula (ver sección I.1 Acotación del objeto de estudio). Por consiguiente, los peligros de la exposición de la privacidad y mal uso aumentan en los contextos educativos en los que se usa *Learning Analytics*. Esto se hace más intrusivo y entraña más peligros cuando se usan de sus derivados *Multimodal Learning Analytics* (Blikstein & Worsley, 2016; Read, 2006) y *Social Network Analytics* (ver sección II.4.3 *Social Network Analytics*) que amplían el campo de recolección de datos personales a aspectos de biometría o estado de ánimo en las interacciones entre estudiantes. Estos métodos suman un peldaño más en la sensibilidad del problema.

El contexto anterior y todas las aproximaciones analíticas mostradas dibujan un contexto de abuso ante la captura y tratamiento de datos educativos, demostrándose así que existe un problema claro de fragilidad en el tratamiento de datos de carácter personal de los estudiantes, su confidencialidad, su seguridad y la protección de su identidad en procesos de *Learning Analytics* en EVA (ver sección I.1 Acotación del objeto de estudio).

Para encontrar al menos una solución al problema y resolver todas las incógnitas que plantea, a continuación, se presentan una serie de resultados relacionados

directamente con los dos objetivos generales, los objetivos específicos y las preguntas de investigación.

En referencia al objetivo específico OE.1.1 (Comprender la necesidad de utilizar una tecnología protectora de la privacidad e identidad de los estudiantes), queda resuelto en la exposición del marco teórico, en el que se detalla y profundiza en cada aspecto del uso y análisis de los datos educativos generados en EVA. En las distintas secciones del marco teórico se discurre desde el origen de Internet, y como su marco tecnológico permite escalar hacia una captura fácil e indiscriminada de grandes cantidades de datos educativos (Williamson, 2017a), hasta la alerta y concienciación al respecto de distintos actores de gran calado en el movimiento de *Learning Analytics* (Lang, Macfadyen, Slade, Prinsloo, & Sclater, 2018; Pardo & Siemens, 2014) .

El segundo objetivo específico OE.1.2 (Validar el uso de la tecnología *blockchain* como protectora de la privacidad e identidad de los estudiantes en procesos de *Learning Analytics*), requiere de un análisis en profundidad que discurre a lo largo de toda la revisión sistemática de la literatura (ver sección II.8 Revisión sistemática de la literatura). Las respuestas a las preguntas de investigación de la SLR demuestran una fuerte inestabilidad e incapacidad de las propuestas que implementan tecnología *blockchain* para ofrecer soluciones robustas y sostenibles. Las razones son múltiples y relacionadas con una inmadurez, inseguridad e ilegalidad tecnológica (ver sección II.8.4.4 Fragilidad). Su inmadurez crea sobre expectativas que en el ámbito educativo hace que aún deba esperarse para obtener resultados sólidos en cuanto a problemas de privacidad en datos almacenados fuera de la cadena. La baja seguridad ante ataques de las distintas vulnerabilidades despierta una desconfianza que hacen preferible la implementación de la tecnología *blockchain* de acceso privado, hecho que se equipara al uso de bases de datos relacionales o de eventos. En cuestiones de ilegalidad, solo algunas soluciones que usan tecnología *blockchain* creadas después de la promulgación del RGPD pueden, de forma muy específica y teórica cumplir con la legalidad. Se concluye para este objetivo que la fragilidad debe resolverse a nivel de bases de datos fuera de la implementación de la tecnología *blockchain*, tecnología ya utilizada en los EVA que incurren en procesos de *Learning Analytics*. Por todo lo expuesto en este párrafo, se descarta la validación del

objetivo y, a su vez, con relación al objetivo general O.1 (Demostrar que el *Blockchain* es una tecnología que puede aportar una posible solución viable al problema de la falta de privacidad, confidencialidad y seguridad de los datos recolectados y usados en procesos de *Learning Analytics* en ecosistemas tecnológicos de aprendizaje), queda descartado el uso de *blockchain* como solución viable. En este sentido, no se desarrolla ningún prototipo basado en *blockchain*, descartando los objetivos específicos OE.1.3 y OE.1.4, y se avanza hacia el objetivo general O.2 y hacia la consecución de sus objetivos específicos asociados. No obstante, se ha experimentado también en el desarrollo de una plataforma basada en tecnología *blockchain* de acceso privado y con permisos que permitiría la gestión, identificación de roles y permisos de acceso a datos personales (ver sección III.5.4 Interoperabilidad con *blockchain*) (Rodríguez, 2019). Se concluye que una implementación de la tecnología *blockchain* puede funcionar como puerta de acceso a datos y facilitar la interoperabilidad entre herramientas, pero no almacenar datos personales de estudiantes y ser una solución al problema de la investigación.

Con respecto al objetivo O.2, y en relación con los objetivos específicos OE.2.1 (Estudiar el impacto del Reglamento General de Protección de Datos en la aplicación de *Learning Analytics*) y OE.2.2 (Estudiar el impacto del Reglamento General de Protección de Datos en los ecosistemas tecnológicos de aprendizaje), se expone que, en el tratamiento de datos de estudiantes, incluidos sus datos personales, se ven implicadas leyes de protección de datos personales. Las leyes internacionales y transferencia de datos entre continentes dibujan un contexto mucho más sensible y complejo, puesto que los acuerdos no regulan el uso que se da de los datos recolectados en el país de destino fuera de la Unión Europea.

Surgen muchas dudas, muchos temores y muchas incógnitas en el tratamiento de datos que un único proyecto o iniciativas, como SHEILA o DELICATE (Drachsler & Greller, 2016; Tsai et al., 2018), no pueden resolver, y menos con políticas o marcos de comportamiento ético (ver sección II.4.4 Miedos y recelos: una cuestión delicada). Por añadidura, España vive convulsa en un cambio continuo de leyes de protecciones de datos. Desde la LORTAD, pasando por la LOPD y terminando en la LOPDGDD y el RGPD, se define un marco de protección de datos para aquellas personas sujetas a términos y

condiciones que les vinculan con un servicio de tratamiento de datos personales. En algunos casos las leyes incluso pueden suponer una desventaja para los usuarios finales, como es el caso de la ley de *cookies* (Zimmerman, 2001).

Se demuestra (ver sección II.6.3 Conocimiento de las leyes educativas), tras entrevistar a más de 250 educadores de España que el desconocimiento de las leyes ha incrementado, quizás debido a las recurrentes renovaciones legales. Al mismo tiempo, y en gran medida, las instituciones educativas hacen un esfuerzo en transferir los conocimientos necesarios para aplicar las leyes en el aula, probablemente mediante mecanismos burocráticos estandarizados y no tanto de conocimientos de la ley. Por contradictorio que parezca, hay más roles que leen políticas de herramientas educativas, pero hay menos preocupación en cuanto a la privacidad y seguridad de los datos de los estudiantes. Es necesaria una concienciación en cuanto a los peligros que supone desproveer a los estudiantes de una privacidad y seguridad adecuada ante los posibles peligros del uso de herramientas educativas digitales y conectadas. Se dan los pasos necesarios para transferir el conocimiento extraído a lo largo de la investigación a distintos profesores mediante la conducción de talleres en dos eventos distintos IBTAC19 (IBTAC, 2019) y *JustKeynote* ("Just Keynote," 2019).

En consideración a los objetivos específicos OE.2.3 (Detectar carencias en los ecosistemas tecnológicos de aprendizaje para la correcta implantación del Reglamento General de Protección de Datos) y OE.2.4 (Desarrollo de un prototipo en un Entorno Virtual de Aprendizaje, como un de los componentes principales de un ecosistema tecnológico de aprendizaje, que asegure un nivel adecuado de confidencialidad y seguridad de datos personales), se realizan entrevistas y se mantienen conversaciones con agentes interesados en el cumplimiento de la legislación de protección de datos personales en EVA (ver sección III.2 Excepciones del RGPD en Moodle). Se decide usar Moodle como plataforma de prototipado, al ser esta la más utilizada en España, e incluso a nivel europeo (Hill, 2016). De las carencias detectadas, se desarrollan paralelamente distintos *plugins* que solucionan aspectos concretos para que Moodle cumpla con el RGPD, y haciendo foco en el anonimato del estudiante usando alias. Desde

la perspectiva de la legalidad, es válido usar un alias para proteger el anonimato de un usuario en un EVA.

Se realiza una serie de entrevistas a distintos perfiles de usuarios de EVA para percibir el nivel de confianza hacia la solución basada en alias. Tras el análisis de las entrevistas, se concluye que asignar un alias a un usuario de un EVA genera suficiente confianza como para desarrollar un prototipo funcional, cumpliéndose así el objetivo general O.2 (Diseñar e implementar una solución tecnológica para adecuar el nivel de confidencialidad de datos personales educativos a lo impuesto por el Reglamento General de Protección de Datos (RGPD) cuando se tienen procesos de *Learning Analytics* en ecosistemas tecnológicos de aprendizaje) mediante el desarrollo de un prototipo funcional expuesto en la sección III.4 Desarrollo de una solución en formato *plugin*. La solución desarrollada en formato *plugin* tiene un impacto elevado, aspecto que se confirma al ser aceptado para su presentación en la conferencia *MoodleMoot Global 2019* (Alier & Amo, 2019). Esta conferencia internacional es de gran importancia puesto que es el encuentro de la comunidad Moodle a escala global, donde estarán presentes actores internacionales de gran calado, agentes intermediarios, organizaciones y fundaciones educativas, desarrolladores y los propios creadores de Moodle.

Paralelamente se realizan distintos esfuerzos de desarrollo generándose otros *plugins*, unos para asegurar un nivel de seguridad en el almacenamiento de los datos en condiciones de confianza y otro en cuando la confianza es muy baja o inexistente (ver sección III.5 Líneas trabajadas en paralelo). Junto a estos esfuerzos aparecen distintas líneas de trabajo que marcan las líneas de futuro.

IV.2. Trabajos futuros

La investigación recorre distintas etapas ordenadas por los dos objetivos generales de esta tesis doctoral. Con respecto al primer objetivo, se descarta implementar una solución con tecnología *blockchain* como línea futura. En cambio, en relación con el segundo objetivo se abren distintas líneas de desarrollo enfocadas a solucionar problemas concretos de los EVA.

Si bien es cierto que algunas líneas se abandonan por causa de los inconvenientes insalvables y necesidades de socio tecnológico (ver III.5.1 Seguridad en la tabla de usuarios), otras líneas son suficientemente interesantes como para ser continuadas:

- Opciones de anonimato por alias (ver III.4 Desarrollo de una solución en formato *plugin*)
- Seguridad aumentada a nivel de SGBD (ver III.5.3 Seguridad en los datos almacenados)
- Buscar socios tecnológicos de las propuestas presentadas

Ambas líneas claras de futuro comportan una nueva revisión del estado de la cuestión, una atención a las potenciales propuestas funcionales y un trabajo de desarrollo que permita extender cada línea en futuras tesis y proyectos. El impacto en la comunidad educativa es un factor que se considera fundamental en la elección de una de las dos líneas. Se considera que el impacto será superior en la propuesta del anonimato, puesto que en una primera instancia alivia problemas graves personales y además valida el objetivo O.2 de la tesis doctoral. El camino abierto en esta investigación es de largo alcance y ocupa un espacio especial en la sociedad.

IV.3. Publicaciones y conferencias a lo largo de la elaboración de la tesis

Se presentan las distintas publicaciones en revistas y conferencias en el ámbito la presente tesis doctoral, relacionadas estas, principalmente con la educación, el *Learning Analytics* y el *blockchain* a lo largo de la elaboración de la tesis (Alier, Amo, García-Peñalvo, Escudero, & Casañ, 2018; Amo, Alier, & Casan, 2018; Amo, Alier, García-Peñalvo, Escudero, et al., 2019; Amo, Alier, García Peñalvo, Fonseca Escudero, & Casañ, 2018; Amo, Fonseca, Alier, García-Peñalvo, Casañ, et al., 2019; Amo, García-Peñalvo, Alier, Escudero, & Casañ, 2018; Campanyà et al., 2019).

La Figura 44 muestra un resumen de las más de 25 contribuciones fruto de la investigación.



Figura 44 Resumen de contribuciones y medios. Elaboración: propia.

IV.3.1. Talleres

- Taller “Clickstream para *Learning Analytics* en Moodle”. MoodleMoot Barcelona, 2018.
- Charla “La medición del proceso de aprendizaje en el aula” in 5th Santillana Compartir *National Congress*. Riviera Maya (México), 2019.
- Taller “Analíticas del aprendizaje in 2° Congreso Internacional de *Flipped* y Metodologías Activas del Aprendizaje”. Misiones (Argentina), 2019.
- Mesa redonda de clausura #EduAnalyticsMOOC en el Curso *Learning Analytics* en Educación (#EduAnalyticsMOOC) de INTEF (Instituto Nacional de Tecnologías Educativas y de la Formación del Profesorado). En línea, 2019. <http://bit.ly/35NqNG2>
- Charla “*Blockchain*: más sombras que luces”. La Salle Talks, Barcelona (España), 2019.

IV.3.2. Congresos

1. Amo-Filvà, D., García-Peñalvo, F. J., Alier Forment, M., Fonseca Escudero, D., & Casañ, M. J. (2018). Privacy and identity management in learning analytics processes with blockchain. *TEEM'18 Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality (Salamanca, Spain, October 24th-26th, 2018)* (pp. 997-1003). New York, NY, USA: ACM. doi:<https://doi.org/10.1145/3284179.3284354>
2. Amo, D., Fonseca, D., Alier, M., García-Peñalvo, F. J., Casañ, M. J., & Alsina, M. (2019). Personal Data Broker: A Solution to Assure Data Privacy in EdTech. In P. Zaphiris & A. Ioannou (Eds.), *Learning and Collaboration Technologies. Design, Experiences. 6th International Conference, LCT 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019. Proceedings, Part I* (pp. 3-14). Cham, Switzerland: Springer Nature. doi: https://doi.org/10.1007/978-3-030-21814-0_1
3. Amo, D., Fonseca, D., Alier, M., García-Peñalvo, F. J., & Casañ, M. J. (2019). Personal Data Broker Instead of Blockchain for Students' Data Privacy Assurance. In Á. Rocha, H. Adeli, L. P. Reis, & S. Costanzo (Eds.), *Knowledge in Information Systems and Technologies* (Vol. 3, pp. 371-380). Switzerland: Springer Nature. doi: https://doi.org/10.1007/978-3-030-16187-3_36
4. Amo, D., Alier, M., Fonseca, D., García-Peñalvo, F. J., Casañ, M. J., & Navarro, J. (2019). Evaluación de la importancia de la ética, privacidad y seguridad en los estudios de Learning Analytics, en el marco de las conferencias LAK. In M. L. Sein-Echaluce Laclea, Á. Fidalgo-Blanco, & F. J. García-Peñalvo (Eds.), *Actas del V Congreso Internacional sobre Aprendizaje, Innovación y Competitividad. CINAIC 2019 (9-11 de Octubre de 2019, Madrid, España)* (pp. 343-348). Zaragoza, Spain: Servicio de Publicaciones Universidad de Zaragoza.
5. Amo, D., Alier, M., García-Peñalvo, F. J., Fonseca, D., & Casañ, M. J. (2019). GDPR Security and Confidentiality compliance in LMS a problem analysis and engineering solution proposal. In M. Á. Conde-González, F. J. Rodríguez-Sedano, C. Fernández-Llamas, & F. J. García-Peñalvo (Eds.), *TEEM'19*

- Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality (Leon, Spain, October 16th-18th, 2019)* (pp. 253-259). New York, NY, USA: ACM.
6. Alier, M., Amo, D. (2019). Modding Moodle to improve confidentiality and privacy support. MoodleMoot Global 2019.
 7. Amo, D., Alier, M., García-Peñalvo, F. J., Fonseca, D., & Casañ, M. J. (2018). Learning Analytics to Assess Students' Behavior With Scratch Through Clickstream. In M. Á. Conde, C. Fernández-Llamas, Á. M. Guerrero-Higueras, F. J. Rodríguez-Sedano, Á. Hernández-García, & F. J. García-Peñalvo (Eds.), *Proceedings of the Learning Analytics Summer Institute Spain 2018 – LASI-SPAIN 2018, (León, Spain, June 18-19, 2018)* (pp. 74-82). Aachen, Germany: CEUR-WS.org.
 8. Amo, D., García-Peñalvo, F. J., Alier, M., Escudero, D. F., & Casañ, M. J. (2018). Privacy and identity management in Learning Analytics processes with Blockchain. *TEEM'18 Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality (Salamanca, Spain, October 24th-26th, 2018)*, (pp. 997–1003). <https://doi.org/10.1145/3284179.3284354>
 9. Amo, D., Fonseca, D., Alier, M., García-Peñalvo, F. J., & Casañ, M. J. (2019). Personal data broker instead of blockchain for students' data privacy assurance. In *Advances in Intelligent Systems and Computing* (Vol. 932, pp. 371–380). https://doi.org/10.1007/978-3-030-16187-3_36
 10. Amo, D., Alier, M., García-Peñalvo, F. J., Fonseca, D., & Casañ, M. J. (2019). GDPR Security and Confidentiality compliance in LMS' a problem analysis and engineering solution proposal. In M. Á. Conde, F. J. Rodríguez, C. Fernández, & F. J. García-Peñalvo (Eds.), *Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM 2019)* (pp. 253–259). New York, NY, USA: ACM.

IV.3.3. Libros y capítulos de libro

- Amo, D., & Santiago, R. (2017). Learning Analytics: la narración del aprendizaje a través de los datos. In Editorial UOC. Barcelona, España: UOC.
- Amo, D. (2018). Analítica del aprendizaje: 30 experiencias con datos en el aula (Edulíticas; D. Amo, ed.). Badalona, España: Amo, Daniel.
- Campanyà, C., Fonseca, D., Martí, N., Amo, D., & Simón, D. (2019). Assessing the Pilot Implementation of Flipped Classroom Methodology in the Concrete and Steel Structures Subject of Architecture Undergraduate Studies. In M. Sein-Echaluce, Á. Fidalgo-Blanco, & F. García-Peñalvo (Eds.), *Innovative Trends in Flipped Teaching and Adaptive Learning* (pp. 55-76). Hershey, PA: IGI Global. doi:10.4018/978-1-5225-8142-0.ch004

IV.3.4. Revistas

- Amo, D., Alier, M., & Casan, M. J. (2018). The student's progress snapshot a hybrid text and visual learning analytics dashboard. *International Journal of Engineering Education*, 34(3), 990–1000. Journal Impact Factor 0,611, Quartile 4.
- Filvà, D. A., Forment, M. A., García-Peñalvo, F. J., Escudero, D. F., & Casañ, M. J. (2019). Clickstream for learning analytics to assess students' behavior with Scratch. *Future Generation Computer Systems*, 93, 673–686. Journal Impact Factor 5,768, Quartile 1. <https://doi.org/10.1016/j.future.2018.10.057>
- Vázquez-Ingelmo, A., García-Peñalvo, F. J., Therón, R., Amo-Filvà, D., & Fonseca-Escudero, D. (2020). Connecting domain-specific features to source code: Towards the automatization of dashboard generation. *Cluster Computing. The Journal of Networks, Software Tools and Applications*, In Press doi:10.1007/s10586-019-03012-1

Apéndices

Referencias detalladas de la SLR

La Tabla 46 enumera las 30 referencias resultado de aplicar la revisión sistemática de la literatura en el campo de educación y *blockchain* ([1]-[30]). Se pretende facilitar la lectura en los distintos apartados y tablas en los que aparecen tales referencias.

Tabla 46 Resumen de referencias encontradas en la revisión sistemática de la literatura

Nº	Título	Autores	Año	Cita
[1]	On the Security and Performance of Proof of Work Blockchains	Arthur Gervais, Ghassan O. Karame, Karl Wüst, Vasileios Glykantzis, Hubert Ritzdorf, Srdjan Capkun	2016	(Gervais et al., 2016)
[2]	Towards Blockchain-enabled School Information Hub	Nelson Bore, Samuel Karumba, Juliet Mutahi, Shelby Solomon Darnell, Charity Wayua, Komminist Weldemariam	2017	(Bore et al., 2017)
[3]	Education Application of Blockchain Technology: Learning Outcome and Meta-Diploma	B. Duan, Y. Zhong, D. Liu	2017	(Duan et al., 2018)
[4]	Hijacking Bitcoin: Routing Attacks on Cryptocurrencies	M. Apostolaki, A. Zohar, L. Vanbever	2017	(Apostolaki et al., 2017)
[5]	ECBC: A High Performance Educational Certificate Blockchain with Efficient Query	Yuqin Xu, Shangli Zhao, Lanju Kong, Yongqing Zheng, Shidong Zhang, Qingzhong Li	2017	(Xu et al., 2017)
[6]	Seeing the Limitations	Daniel Drescher	2017	(Drescher, 2017)

[7]	From Bitcoin to Cybersecurity: a Comparative Study of Blockchain Application and Security Issues	Dai, Fangfang; Shi, Yue; Meng, Nan; Wei, Liang; Ye, Zhiguo	2017	(F. Dai et al., 2017)
[8]	Towards a new Ubiquitous Learning Environment Based on Blockchain Technology	Bdiwi, Rawia; de Runz, Cyril; Faiz, Sami; Cherif, Arab Ali	2017	(Bdiwi et al., 2017)
[9]	A Novel Blockchain-based Education Records Verification Solution	Meng Han, Zhigang Li and Jing (Selena) He, Dalei Wu, Ying Xie, Asif Baba	2018	(Han et al., 2018)
[10]	Connecting Decentralized Learning Records: A Blockchain Based Learning Analytics Platform	Patrick Ocheja, Brendan Flanagan, Hiroaki Ogata	2018	(Ocheja et al., 2018)
[11]	Security Vulnerabilities in Ethereum Smart Contracts	Alexander Mense, Markus Flatscher	2018	(Mense & Flatscher, 2018)
[12]	General Data Protection Regulation Complied Blockchain Architecture for Personally Identifiable Information Management	N. Al-Zaben; M. M. Hassan Onik; J. Yang; N. Lee; C. Kim	2018	(Al-Zaben et al., 2019)
[13]	A Blueprint for a Blockchain-Based Architecture to Power a Distributed Network of Tamper-Evident	J. C. Farah; A. Vozniuk; M. J. Rodríguez-Triana; D. Gillet	2018	(Farah et al., 2018)

	Learning Trace Repositories			
[14]	Parallel-Education- Blockchain Driven Smart Education: Challenges and Issues	X. Gong; X. Liu; S. Jing; G. Xiong; J. Zhou	2018	(Gong et al., 2019)
[15]	EduCTX: A Blockchain- Based Higher Education Credit Platform	M. Turkanović; M. Hölbl; K. Košič; M. Heričko; A. Kamišalić	2018	(Turkanović et al., 2018)
[16]	CredenceLedger: A Permissioned Blockchain for Verifiable Academic Credentials	R. Arenas; P. Fernández	2018	(Arenas & Fernandez, 2018)
[17]	A Distributed Credit Transfer Educational Framework based on Blockchain	A. Srivastava; P. Bhattacharya; A. Singh; A. Mathur; O. Prakash; R. Pradhan	2018	(Srivastava et al., 2019)
[18]	Blockchain for Student Data Privacy and Consent	S. Gilda; M. Mehrotra	2018	(Gilda & Mehrotra, 2018)
[19]	Scholarium: Supporting Identity Claims Through a Permissioned Blockchain	G. Dima; A. Jitariu; C. Pisa; G. Bianchi	2018	(Dima et al., 2018)
[20]	A Survey on Security and Privacy Issues of Bitcoin	M. Conti; E. Sandeep Kumar; C. Lal; S. Ruj	2018	(Conti et al., 2018)
[21]	On blockchain security and relevant attacks	J. Moubarak; E. Filiol; M. Chamoun	2018	(Moubarak et al., 2018)

[22]	A Survey of Attacks on the Bitcoin System	A. Soni; S. Maheshwari	2018	(Soni & Maheshwari, 2018)
[23]	Exploring blockchain technology and its potential applications for education	Guang Chen, Bing Xu, Manli Lu, Nian-Shing Chen	2018	(G. Chen et al., 2018)
[24]	Learning analytics platform in higher education in Japan	Flanagan, Brendan; Ogata, Hiroaki	2018	(Flanagan & Ogata, 2018)
[25]	Blockchain and law: Incompatible codes?	Millard, Christopher	2018	(Millard, 2018)
[26]	A survey on security and privacy issues of blockchain technology	Joshi, Archana Prashanth; Han, Meng; Wang, Yan	2018	(Joshi et al., 2018)
[27]	Blockchain and its Potential in Education	Turcu, Cristina; Turcu, Cornel; Chiuchisan, Iuliana	2018	(Turcu et al., 2018)
[28]	Application of Blockchain Technology in Online Education	Sun, Han; Wang, Xiaoyue; Wang, Xinge	2018	(Sun et al., 2018)
[29]	Chronicle of a Clash Foretold: Blockchains and the GDPR's Right to Erasure	Pagallo, Ugo; Bassi, Eleonora; Crepaldi, Marco; Durante, Massimo	2018	(Pagallo et al., 2018)
[30]	Blockchain Demystified: A Technical and Legal Introduction to Distributed and Centralised Ledgers	J. Bacon, J. Michels, C. Millard et al.	2018	(Bacon et al., 2018)

Localizaciones de las respuestas a preguntas y consecución de objetivos

Las siguientes tablas localizan a lo largo del manuscrito la respuesta a las preguntas de la investigación (PI1-7) y la consecución de los objetivos.

Tabla 47 Puntos del manuscrito en los que se responden las preguntas de investigación.

Pregunta de investigación	Respuesta
PI1. ¿Cuáles son los orígenes y evolución del análisis de datos educativos?	Su respuesta se expone en el Marco teórico, en las secciones II.3 MOOC y II.4 <i>Educational Data Mining</i>
PI2. ¿Cuáles son los problemas relacionados con el análisis de datos educativos?	Su respuesta se expone en el Marco teórico, en las secciones II.4.4 Miedos y recelos: una cuestión delicada y II.5 <i>Learning Analytics y Conocimiento en congresos</i>
PI3. ¿Qué implicaciones tiene asegurar la privacidad y seguridad de los datos educativos en procesos de <i>Learning Analytics</i> ?	Su respuesta se expone en el Marco teórico, en la sección II.6 Leyes sobre protección de datos personales
PI4. ¿Cuáles son las soluciones que implementan la tecnología <i>blockchain</i> para resolver problemas en el contexto educativo?	Su respuesta se expone en el Marco teórico, en la sección II.8.3.1 RQ1. ¿Qué soluciones se han aportado en el campo de estudio? Y II.8.4.1 Soluciones educativas, como parte de la Revisión sistemática de la literatura
PI5. ¿Existe alguna aplicación concreta de la tecnología <i>blockchain</i> que pueda considerarse como una solución al problema planteado y cumplir con las leyes de protección de datos?	Su respuesta se expone en el Marco teórico, en la sección II.8.3.3 RQ3. ¿La tecnología <i>blockchain</i> cumple con el RGPD? y II.8.4.3 Legalidad, como parte de la Revisión sistemática de la literatura
PI6. ¿Puede solucionarse el problema con una implementación de la tecnología	Su respuesta se expone en el Marco teórico, en la sección II.8.3.4 RQ4. ¿Qué puede resolver la tecnología <i>blockchain</i> en relación

<i>blockchain</i> sin la ayuda de otras tecnologías de almacenamiento de datos?	con el problema? Y II.8.4.4 Fragilidad, como parte de la Revisión sistemática de la literatura
PI7. ¿Puede asegurarse un adecuado nivel de protección de datos en entornos de aprendizaje sin necesidad de usar tecnología <i>blockchain</i> ?	Su respuesta afirmativa se expone en el Marco empírico, a lo largo del desarrollo y exposición de los distintos prototipos funcionales en las secciones III.4 Desarrollo de una solución en formato <i>plugin</i> y III.5 Líneas trabajadas en paralelo, como parte de la Fase 2 del O.2

Tabla 48 Puntos del manuscrito en los que se expone la consecución de los objetivos específicos del Objetivos 1

Objetivos Específicos del O. 1	Consecución
OE.1.1. Comprender la necesidad de utilizar una tecnología protectora de la privacidad e identidad de los estudiantes	Su consecución positiva se expone en el Marco teórico, a lo largo de las secciones II.5 <i>Learning Analytics</i> y II.6.3 Conocimiento de las leyes educativas
OE.1.2. Validar la tecnología <i>blockchain</i> como protectora de la privacidad e identidad de los estudiantes en procesos de <i>Learning Analytics</i>	No se puede conseguir ni validar tras el análisis de los resultados de la SLR expuesta en el Marco teórico, en la sección II.8 Revisión sistemática de la literatura
OE.1.3. Desarrollo de un prototipo que implemente la tecnología <i>blockchain</i> para proteger la privacidad e identidad de los estudiantes en el uso de <i>Learning Analytics</i>	No se considera un posible desarrollo tras el análisis de los resultados de la SLR expuesta en el Marco teórico, en la sección II.8 Revisión sistemática de la literatura

Tabla 49 Puntos del manuscrito en los que se expone la consecución de los objetivos específicos del Objetivos 2

Objetivos Específicos del O.2	Consecución
OE.2.1. Estudiar el impacto del Reglamento General de Protección de Datos en la aplicación de <i>Learning Analytics</i>	Su consecución positiva se expone en el Marco teórico, a lo largo de las secciones II.5 <i>Learning Analytics</i> y Conocimiento en congresos, II.6 Leyes sobre protección de

	datos personales y II.6.3 Conocimiento de las leyes educativas
OE.2.2. Estudiar el impacto del Reglamento General de Protección de Datos en los ecosistemas tecnológicos de aprendizaje	Su consecución positiva se expone en el Marco teórico, a lo largo de las secciones II.5 <i>Learning Analytics</i> y <i>Conocimiento en congresos Learning Analytics</i> y , II.6 Leyes sobre protección de datos personales y III.2 Excepciones del RGPD en Moodle
OE2.3. Detectar carencias en los ecosistemas tecnológicos de aprendizaje para la correcta implantación del Reglamento General de Protección de Datos	Su consecución positiva se expone en el Marco empírico, a lo largo de las secciones III.2 Excepciones del RGPD en Moodle y III.3 Experiencia de usuario
OE2.4. Desarrollo de un prototipo en un Entorno Virtual de Aprendizaje, como un de los componentes principales de un ecosistema tecnológico de aprendizaje, que asegure un nivel adecuado de confidencialidad y seguridad de datos personales del estudiante impuesto por el Reglamento General de Protección de Datos	Su consecución positiva se expone en el Marco empírico, a lo largo del desarrollo y exposición de los distintos prototipos funcionales en las secciones III.4 Desarrollo de una solución en formato <i>plugin</i> y III.5 Líneas trabajadas en paralelo

Referencias

- 3iPunt. (2019). Retrieved February 1, 2018, from <https://www.tresipunt.com/>
- Adell, J., Bartolomé, A., Bellver, C., Domingue, J., Dos Santos, A., Guimarães, C., ... Watters, A. (2018). Blockchain en educación, cadenas rompiendo moldes – Institut de Recerca en Educació. In A. Bartolomé & J. M. Manuel-Ferrer (Eds.), *Colección Transmedia XXI* (LMI). Retrieved from <http://www.ub.edu/ire/en/new-book-blockchain-en-educacion-cadenas-rompiendo-moldes/>
- Adell, J., Bellver, A. J., & Bellver, C. (2008). Entornos virtuales de aprendizaje y estándares de e-learning. *Psicología de La Educación Virtual. Enseñar y Aprender Con Las Tecnologías de La Información y La Comunicación*, 274–298.
- Adkins, D. (2009). The use of social network analysis to measure knowledge sharing in the New York State Project Management Community of Practice. (University of Albany, State University of New York; Vol. 69). Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2009-99010-247&site=ehost-live>
- AGPD. (2016). *Guía Acerca Del Escudo De Privacidad UE - EE. UU.* 24. <https://doi.org/10.2838/12912>
- Al-Samarraie, H., & Saeed, N. (2018). A systematic review of cloud computing tools for collaborative learning: Opportunities and challenges to the blended-learning environment. *Computers & Education*, 124, 77–91. <https://doi.org/https://doi.org/10.1016/j.compedu.2018.05.016>
- Al-Zaben, N., Onik, M. M. H., Yang, J., Lee, N. Y., & Kim, C. S. (2019). General Data Protection Regulation Complied Blockchain Architecture for Personally Identifiable Information Management. *Proceedings - 2018 International Conference on Computing, Electronics and Communications Engineering, ICCECE 2018*, 77–82. <https://doi.org/10.1109/iCCECOME.2018.8658586>
- Alier, M., & Amo, D. (2019). Modding Moodle to improve confidentiality and privacy support. 2:10pm (PROGRAM – DAY 2). Retrieved from MoodleMoot Global 2019 website: <https://moodleMoot.org/mootglobal19/day-2/>

- Alier, M., Amo, D., García-Peñalvo, F. J., Escudero, D. F., & Casañ, M. J. (2018). Learning analytics' privacy on the blockchain. *ACM International Conference Proceeding Series*, 294–298. <https://doi.org/10.1145/3284179.3284231>
- Alier, M., & Casañ, M. J. (2008). Creating wiki communities in blended learning environment and the creation of the Moodle New Wiki. *Proceedings of the 2008 International Conference on Frontiers in Education: Computer Science and Computer Engineering, FECS 2008*, 1(3), 99–105.
- Alier, M., Casañ, M. J., Conde, M. Á., García-Penalvo, F. J., & Severance, C. (2010). Interoperability for LMS: the missing piece to become the common place for e-learning innovation. *International Journal of Knowledge and Learning (IJKL)*, 6(2–3), 130–141. <https://doi.org/10.1504/IJKL.2010.034749>
- Almatrafi, O., & Johri, A. (2018). Systematic Review of Discussion Forums in Massive Open Online Courses (MOOCs). *IEEE Transactions on Learning Technologies*. <https://doi.org/10.1109/TLT.2018.2859304>
- Amo, D. (2018). Tesis Daniel Amo Filvà - Revisión Sistemática Literatura. Retrieved from https://lasalleuniversities-my.sharepoint.com/:x:/g/personal/daniel_amo_salle_url_edu/EQoQGKr7zuBlpYkA0gB4mGoBEP7NCufbBjhkURpKOAklyQ?e=0cgvVO
- Amo, D., & Alier, M. (2019). Moodle Personal Data Brocker Log Store. Retrieved May 14, 2019, from GitHub website: https://github.com/danielamof/logstore_pdb
- Amo, D., Alier, M., & Casan, M. J. (2018). The student's progress snapshot a hybrid text and visual learning analytics dashboard. *International Journal of Engineering Education*, 34(3), 990–1000.
- Amo, D., Alier, M., García-Peñalvo, F. J., Escudero, D. F., & Casañ, M. J. (2019). Clickstream for learning analytics to assess students' behavior with Scratch. *Future Generation Computer Systems*, 93, 673–686. <https://doi.org/10.1016/j.future.2018.10.057>
- Amo, D., Alier, M., García-Peñalvo, F. J., Fonseca, D., & Casañ, M. J. (2019). GDPR Security and Confidentiality compliance in LMS' a problem analysis and engineering solution

- proposal. In Miguel Ángel Conde, F. J. Rodríguez, C. Fernández, & F. J. García-Peñalvo (Eds.), *Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM 2019)* (pp. 253–259). New York, NY, USA: ACM.
- Amo, D., Alier, M., García Peñalvo, F. J., Fonseca Escudero, D., & Casañ, M. J. (2018). Learning analytics to assess students' behavior with scratch through clickstream. *Proceedings of the Learning Analytics Summer Institute Spain 2018: León, Spain, June 18-19, 2018*, 74–82.
- Amo, D., Casañ, M. J., & Alier, M. (2014). Google analytics for time behavior measurement in Moodle. *Iberian Conference on Information Systems and Technologies, CISTI*, 1–6. <https://doi.org/10.1109/CISTI.2014.6877095>
- Amo, D., Fonseca, D., Alier, M., García-Peñalvo, F. J., & Casañ, M. J. (2019). Personal data broker instead of blockchain for students' data privacy assurance. In Á. Rocha, H. Adeli, L. P. Reis, & S. Costanzo (Eds.), *Knowledge in Information Systems and Technologies* (Vol. 3). https://doi.org/10.1007/978-3-030-16187-3_36
- Amo, D., Fonseca, D., Alier, M., García-Peñalvo, F. J., Casañ, M. J., & Alsina, M. (2019). Personal Data Broker: A Solution to Assure Data Privacy in EdTech. *International Conference on Human-Computer Interaction*, 3–14. https://doi.org/10.1007/978-3-030-21814-0_1
- Amo, D., García-Peñalvo, F. J., & Alier, M. (2014). Social network analysis approaches for social learning support. In F. J. García-Peñalvo (Ed.), *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM 2014)* (pp. 269–274). <https://doi.org/10.1145/2669711.2669910>
- Amo, D., García-Peñalvo, F. J., Alier, M., Escudero, D. F., & Casañ, M. J. (2018). Privacy and identity management in Learning Analytics processes with Blockchain. *TEEM'18 Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality (Salamanca, Spain, October 24th-26th, 2018)*, 997–1003. <https://doi.org/10.1145/3284179.3284354>

- Amo, D., & Santiago, R. (2017). Learning Analytics: la narración del aprendizaje a través de los datos. In *Editorial UOC*. Barcelona, España: UOC.
- Apostolaki, M., Zohar, A., & Vanbever, L. (2017). Hijacking Bitcoin: Routing Attacks on Cryptocurrencies. *Proceedings - IEEE Symposium on Security and Privacy*, 375–392. <https://doi.org/10.1109/SP.2017.29>
- Arenas, R., & Fernandez, P. (2018). CredenceLedger: A Permissioned Blockchain for Verifiable Academic Credentials. *2018 IEEE International Conference on Engineering, Technology and Innovation, ICE/ITMC 2018 - Proceedings*, 1–6. <https://doi.org/10.1109/ICE.2018.8436324>
- Aste, T., Tasca, P., & Di Matteo, T. (2017). Blockchain Technologies: The Foreseeable Impact on Society and Industry. *Computer*, 50(9), 18–28. <https://doi.org/10.1109/MC.2017.3571064>
- Avanza, P. (2005). *Plan Avanza 2*. Madrid.
- Aznarte, J. L., Hidalgo, R., Rubió, E., & Ruipérez, J. A. (2019). Canal UNED - Inteligencia Artificial y la personalización del aprendizaje (Mesa redonda). Retrieved June 27, 2019, from <https://canal.uned.es/video/5caf043ba3eeb095388b4567>
- Bacon, J., Michels, J. D., Millard, C., & Singh, J. (2018). Blockchain Demystified: a technical and legal introduction to distributed and centralised ledgers. *Richmond Journal of Law & Technology*, XXV(1), 1–106. Retrieved from <https://jolt.richmond.edu/files/2018/11/Michelsetal-Final-1.pdf>
- Badge, J., Johnson, S., Moseley, A., & Cann, A. (2011). Observing Emerging Student Networks on a Microblogging Service. *Journal of Online Learning and Teaching*, 7(1), 90. Retrieved from <https://reddog.rmu.edu/login?url=http://search.proquest.com/docview/1497199183?accountid=28365%5Cnhttp://linksource.ebsco.com/linking.aspx?sid=ProQ%3Aeducation&fmt=journal&genre=article&issn=&volume=7&issue=1&date=2011-03-01&spage=90&title=Journal+of+Onli>
- Baggaley, J. (2013). MOOC rampant. *Distance Education*, 34(3), 368–378. <https://doi.org/10.1080/01587919.2013.835768>

- Bakharia, A., Heathcote, E., & Dawson, S. (2009). Social networks adapting pedagogical practice: SNAPP. Citeseer.
- Baricco, A. (2019). *The game*. Einaudi.
- Bartolomé, A., & Lindín, C. (2019). Posibilidades del Blockchain en Educación. *Education in the Knowledge Society (EKS)*, 19(4), 81. <https://doi.org/10.14201/eks20181948193>
- Bartolomé, A., Manuel, J., & Ferrer, M. (2018). Blockchain en Educación - Cadenas rompiendo moldes. *Virtualidad, Educación y Ciencia*, 9(17), 114–116. Retrieved from <http://www.lmi.ub.es/transmedia>
- Bartolomé Pina, A. R., Bellver Torlà, C., Castañeda Quintero, L., & Adell Segura, J. (2017). Blockchain en Educación: introducción y crítica al estado de la cuestión. *Edutec. Revista Electrónica de Tecnología Educativa*, (61), a363. <https://doi.org/10.21556/edutec.2017.61.915>
- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. *Third International AAAI Conference on Weblogs and Social Media*, 361–362. <https://doi.org/10.1136/qshc.2004.010033>
- Baturones, J. J. E. (1998). La regulación de los datos sensibles a la Directiva 95/46/CE. *Informática y Derecho: Revista Iberoamericana de Derecho Informático*, (23), 1217–1248.
- Bdiwi, R., Runz, C. de, Faiz, S., & Cherif, A. A. (2017). Towards a New Ubiquitous Learning Environment Based on Blockchain Technology. *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*, 101–102. <https://doi.org/10.1109/ICALT.2017.37>
- Berendt, B., Littlejohn, A., Kern, P., Mtros, P., Shacklock, X., & Blakemore, M. (2017). *Big data for monitoring educational systems*. <https://doi.org/10.2766/38557>
- Berland, M., Baker, R. S., & Blikstein, P. (2014). Educational data mining and learning analytics: Applications to constructionist research. *Technology, Knowledge and Learning*, 19(1–2), 205–220. <https://doi.org/10.1007/s10758-014-9223-7>
- Berlanga, A. J., Peñalvo, F. J. G., & Sloep, P. B. (2010). Towards eLearning 2.0 University.

- Interactive Learning Environments*, Vol. 18, pp. 199–201.
<https://doi.org/10.1080/10494820.2010.500498>
- Berners-Lee, T. (1992). The world-wide web. *Computer Networks and ISDN Systems*, 25(4–5), 454–459. Retrieved from <https://cds.cern.ch/record/245440/files/p69.pdf>
- Berners-Lee, T., Connolly, D., Muldrow, K., & DTDs, S. (1986). HTML 2.0. *RFC1866*. *Ftp://Ds. Internic. Net/Rfc/Rfc1866. Txt* (Nov. 1995). Retrieved from <ftp://ds.internic.net/rfc/rfc1866.txt>
- Berners-Lee, T., Fielding, R., & Frystyk, H. (1996). *Hypertext transfer protocol--HTTP/1.0*. Retrieved from <http://www.hjp.at/doc/rfc/rfc1945.html>
- Bevan, N., Carter, J., & Harker, S. (2015). Iso 9241-11 revised: What have we learnt about usability since 1998? *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9169, 143–151. https://doi.org/10.1007/978-3-319-20901-2_13
- Bienkowski, M., Feng, M., & Means, B. (2014). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. *Educational Improvement Through Data Mining and Analytics*, 1–60.
- Blikstein, P. (2013). Multimodal learning analytics. *ACM International Conference Proceeding Series*, 102–106. <https://doi.org/10.1145/2460296.2460316>
- Blikstein, P., & Worsley, M. (2016). Multimodal Learning Analytics and Education Data Mining: Using Computational Technologies to Measure Complex Learning Tasks. *Journal of Learning Analytics*, 3(2), 220–238. <https://doi.org/10.18608/jla.2016.32.11>
- Bode, M., & Kristensen, D. B. (2015). The digital doppelgänger within: A study on self-tracking and the quantified self movement. *Assembling Consumption: Researching Actors, Networks and Markets*, 119–134. <https://doi.org/10.4324/9781315743608>
- Boletín Oficial del Estado. (2018). LOPDGDD BOE-A-2018-16673. Retrieved September 6, 2019, from Boletín Oficial del Estado, núm. 294 website: <https://www.boe.es/buscar/doc.php?id=BOE-A-2018-16673>

- Boletín Oficial del Estado. (2019). Documento BOE-A-2019-15790. Retrieved October 31, 2019, from Boletín Oficial del Estado, núm. 266 website: <https://www.boe.es/buscar/doc.php?id=BOE-A-2019-15790>
- Bore, N., Karumba, S., Mutahi, J., Darnell, S. S., Wayua, C., & Weldemariam, K. (2017). Towards Blockchain-enabled school information hub. *ACM International Conference Proceeding Series, Part F1320*, 1–4. <https://doi.org/10.1145/3136560.3136584>
- Borge, M., Kokoris-Kogias, E., Jovanovic, P., Gasser, L., Gailly, N., & Ford, B. (2017). Proof-of-personhood: Redemocratizing permissionless cryptocurrencies. *Proceedings - 2nd IEEE European Symposium on Security and Privacy Workshops, EuroS and PW 2017*, 23–26. <https://doi.org/10.1109/EuroSPW.2017.46>
- Botsman, R. (2012). The currency of the new economy is trust. In *Ted*. Retrieved from <http://on.ted.com/Botsman>
- Brinton, C. G., Buccapatnam, S., Chiang, M., & Poor, H. V. (2016). Mining MOOC Clickstreams: Video-Watching Behavior vs. In-Video Quiz Performance. *IEEE Transactions on Signal Processing*, 64(14), 3677–3692. <https://doi.org/10.1109/TSP.2016.2546228>
- Brinton, C. G., & Chiang, M. (2015). MOOC performance prediction via clickstream data and social learning networks. *Proceedings - IEEE INFOCOM, 26*, 2299–2307. <https://doi.org/10.1109/INFOCOM.2015.7218617>
- Broad, W. J. (1992). Clinton to Promote High Technology. Retrieved September 5, 2019, from *Nyt* website: <https://www.nytimes.com/1992/11/10/science/clinton-to-promote-high-technology-with-gore-in-charge.html>
- Calvard, T. S. (2016). Big data, organizational learning, and sensemaking: Theorizing interpretive challenges under conditions of dynamic complexity. *Management Learning*, 47(1), 65–82. <https://doi.org/10.1177/1350507615592113>
- Campanyà, C., Fonseca, D., Martí, N., Amo, D., & Simón, D. (2019). Assessing the Pilot Implementation of Flipped Classroom Methodology in the Concrete and Steel Structures Subject of Architecture Undergraduate Studies. *Innovative Trends in*

Flipped Teaching and Adaptive Learning, 55–76. <https://doi.org/10.4018/978-1-5225-8142-0.ch004>

Campbell, J. P., DeBlois, P. B., & Oblinger, D. G. (2007). Academic Analytics: A New Tool for a New Era. *Educause Review*, 42(4), 40-42,44,46,48,50,52,54,56-57. Retrieved from

<http://net.educause.edu/ir/library/pdf/ERM0742.pdf>
<http://net.educause.edu/ir/library/pdf/erm0742.pdf>
<http://eric.ed.gov/ERICWebPortal/recordDetail?accno=EJ769402>
<http://net.educause.edu/ir/library/pdf/ERM0742.pdf>
<http://net.educause.edu/ir/li>

Capuano, N., Mangione, G. R., Mazzoni, E., Miranda, S., & Orciuoli, F. (2014). Wiring role taking in collaborative learning environments. SNA and semantic web can improve CSCL script? *International Journal of Emerging Technologies in Learning*, 9(7), 30–38. <https://doi.org/10.3991/ijet.v9i7.3719>

Chan, A. (2014). Forum Graph. Retrieved September 18, 2019, from Moodle plugins directory https://moodle.org/plugins/pluginversions.php?plugin=report_forumgraph website:

Charleer, S., Klerkx, J., Duval, E., De Laet, T., & Verbert, K. (2016). *Creating Effective Learning Analytics Dashboards: Lessons Learnt*. https://doi.org/10.1007/978-3-319-45153-4_4

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, 4(5–6), 318–331. <https://doi.org/10.1504/IJTEL.2012.051815>

Chen, G., Xu, B., Lu, M., & Chen, N.-S. (2018). Exploring blockchain technology and its potential applications for education. *Smart Learning Environments*, 5(1), 1. <https://doi.org/10.1186/s40561-017-0050-x>

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly: Management Information Systems*, 36(4), 1165–1188.

Christensen, C. M., & Bower, J. L. (1995). Disruptive technologies: catching the wave.

- Long Range Planning*, 28(2), 155. [https://doi.org/https://doi.org/10.1016/0024-6301\(95\)91075-1](https://doi.org/https://doi.org/10.1016/0024-6301(95)91075-1)
- Clark, D. (2013). Adaptive MOOCs. *CogBooks Adaptive Learning*. Retrieved from [file:///C:/Users/owner/Downloads/Whhttp_adaptiveMOOCs_CC_Final \(WGN&DC\).pdf](file:///C:/Users/owner/Downloads/Whhttp_adaptiveMOOCs_CC_Final(WGN&DC).pdf)
- Clayton, C. (2015). Disruptive Innovation. https://doi.org/10.1007/978-3-658-24702-7_3
- Cody, J. P. (1998). Protecting privacy over the Internet: Has the time come to abandon self-regulation. *Cath. UL Rev.*, 48, 1183. Retrieved from <https://scholarship.law.edu/cgi/viewcontent.cgi?article=1430&context=lawreview>
- Conde, Miguel Á., García-Peñalvo, F. J., Rodríguez-Conde, M. J., Alier, M., Casañ, M. J., & Piguillem, J. (2014). An evolving Learning Management System for new educational environments using 2.0 tools. *Interactive Learning Environments*, 22(2), 188–204. <https://doi.org/10.1080/10494820.2012.745433>
- Conde, Miguel Á., & Hernández-García, Á. (2015). Learning analytics for educational decision making. *Computers in Human Behavior*, 47(47), 1–3. <https://doi.org/10.1016/j.chb.2014.12.034>
- Conte de Leon, D., Stalick, A. Q., Jillepalli, A. A., Haney, M. A., & Sheldon, F. T. (2017). Blockchain: properties and misconceptions. *Asia Pacific Journal of Innovation and Entrepreneurship*, 11(3), 286–300. <https://doi.org/10.1108/apjie-12-2017-034>
- Conti, M., Sandeep, K. E., Lal, C., & Ruj, S. (2018). A survey on security and privacy issues of bitcoin. *IEEE Communications Surveys and Tutorials*, 20(4), 3416–3452. <https://doi.org/10.1109/COMST.2018.2842460>
- Costa, C., Teixeira, L., & Alvelos, H. (2018). Exploring the usage of MOOCs in higher education institutions: Characterization of the most used platforms. *International Journal of Information and Communication Technology Education*, 14(4), 1–17. <https://doi.org/10.4018/IJICTE.2018100101>
- Cruz-Benito, J., Borrás-Gene, O., García-Peñalvo, F. J., Blanco, A. F., & Theron, R. (2017).

- Learning Communities in Social Networks and Their Relationship with the MOOCs. *Revista Iberoamericana de Tecnologías Del Aprendizaje*, 12(1), 24–36. <https://doi.org/10.1109/RITA.2017.2655218>
- Cruz-Benito, J., García-Peñalvo, F. J., & Therón, R. (2019). Analyzing the software architectures supporting HCI/HMI processes through a systematic review of the literature. *Telematics and Informatics*, 38, 118–132. <https://doi.org/10.1016/j.tele.2018.09.006>
- D’Antoni, S. (2012). The UNESCO OER community 2005-2009: From collective interaction to collaborative action. In A. Okada, T. Connolly, & P. J. Scott (Eds.), *Collaborative Learning 2.0: Open Educational Resources* (pp. 16–37). <https://doi.org/10.4018/978-1-4666-0300-4.ch002>
- Dai, F., Shi, Y., Meng, N., Wei, L., & Ye, Z. (2017). From Bitcoin to cybersecurity: A comparative study of blockchain application and security issues. *2017 4th International Conference on Systems and Informatics (ICSAI)*, 975–979. <https://doi.org/10.1109/ICSAI.2017.8248427>
- Dai, W. (1998). Alternative b-money creation. Retrieved June 1, 2018, from weidai.com website: <http://www.weidai.com/bmoney.txt>
- Daniel, B. K. (2019). Big Data and data science: A critical review of issues for educational research. *British Journal of Educational Technology*, 50(1), 101–113. <https://doi.org/10.1111/bjet.12595>
- Daniel, J. (2012). Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility. *Journal of Interactive Media in Education*, 2012(3), 18. <https://doi.org/10.5334/2012-18>
- Dawson, S. (2008). A study of the relationship between student social networks and sense of community. *Educational Technology and Society*, 11(3), 224–238.
- De Waard, I. (2013). MOOC Yourself - Set up your own MOOC for Business, Non-Profits, and Informal Communities. *Kindle*.
- Del Blanco, A., Serrano, A., Freire, M., Martinez-Ortiz, I., & Fernandez-Manjon, B. (2013). E-Learning standards and learning analytics. Can data collection be improved by

- using standard data models? *IEEE Global Engineering Education Conference, EDUCON*, 1255–1261. <https://doi.org/10.1109/EduCon.2013.6530268>
- Dexter, S. (2018). Blockchain vs DLT. Retrieved September 25, 2019, from mangoresearch.com website: <https://www.mangoresearch.co/blockchain-vs-distributed-ledger-technology-dlt/>
- Dima, G. A., Jitariu, A. G., Pisa, C., & Bianchi, G. (2018). Scholarium: Supporting Identity Claims Through a Permissioned Blockchain. *IEEE 4th International Forum on Research and Technologies for Society and Industry, RTSI 2018 - Proceedings*, 1–6. <https://doi.org/10.1109/RTSI.2018.8548407>
- Downes, S. (2012). The Rise of MOOCs. Retrieved from Knowledge, Learning, Community (Vol. 2015) website: <http://www.downes.ca/post/57911>
- Drachler, H. (2016). Ethics & Privacy in Learning Analytics - a DELICATE issue - LACE - Learning Analytics Community Exchange. Retrieved September 20, 2017, from LACE website: <http://www.laceproject.eu/blog/ethics-privacy-in-learning-analytics-a-delicate-issue/>
- Drachler, H., & Greller, W. (2016). Privacy and analytics: it's a DELICATE issue a checklist for trusted learning analytics. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, 25-29-April, 89–98*. <https://doi.org/10.1145/2883851.2883893>
- Drescher, D. (2017). *Blockchain Basics*. Apress.
- Duan, B., Zhong, Y., & Liu, D. (2018). Education application of blockchain technology: Learning outcome and meta-diploma. *Proceedings of the International Conference on Parallel and Distributed Systems - ICPADS, 2017-Decem, 814–817*. <https://doi.org/10.1109/ICPADS.2017.00114>
- Duval, E. (2012). Learning Analytics and Educational Data Mining. Retrieved July 24, 2018, from Erik Duval's Weblog website: <https://erikduval.wordpress.com/2012/01/30/learning-analytics-and-educational-data-mining/>
- Egelman, S., Felt, A. P., & Wagner, D. (2013). Choice architecture and smartphone

- privacy: There's a price for that. In *The Economics of Information Security and Privacy* (pp. 211–236). https://doi.org/10.1007/978-3-642-39498-0_10
- EP and the CEU. (2016). Regulation (EU) 2016/679 GDPR. Retrieved June 27, 2019, from Official Journal of the European Union website: <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32016R0679>
- Epelboin, Y. (2013). MOOC: a revolution in teaching? A European view. *EUNIS 2013 Congress Proceedings: 2013: ICT Role for Next Generation Universities*, 1(1). <https://doi.org/10.7250/eunis.2013.003>
- Erlin, Yusof, N., & Rahman, A. A. (2008). Integrating content analysis and social network analysis for analyzing asynchronous discussion forum. *Proceedings - International Symposium on Information Technology 2008, ITSIM*, 4, 1–8. <https://doi.org/10.1109/ITSIM.2008.4631996>
- Exodus Privacy. (2019). Exodus Privacy Code of Conduct. Retrieved June 12, 2019, from <https://exodus-privacy.eu.org/en/page/coc/>
- Farah, J. C., Vozniuk, A., Rodriguez-Triana, M. J., & Gillet, D. (2018, July 10). A blueprint for a blockchain-based architecture to power a distributed network of tamper-evident learning trace repositories. *Proceedings - IEEE 18th International Conference on Advanced Learning Technologies, ICALT 2018*, pp. 218–222. <https://doi.org/10.1109/ICALT.2018.00059>
- Farrell, S., & Tschofenig, H. (2014). Pervasive Monitoring Is an Attack. Retrieved from Internet Engineering Task Force website: <https://tools.ietf.org/html/rfc7258>
- Feldstein, M. (2010). Social Network Analysis and the LMS. Retrieved July 1, 2014, from <https://eliterate.us/social-network-analysis-and-the-lms/>
- Fidalgo-Blanco, Á., Martínez-Nuñez, M., Borrás-Gene, O., & Sánchez-Medina, J. J. (2017). Micro flip teaching – An innovative model to promote the active involvement of students. *Computers in Human Behavior*, 72, 713–723. <https://doi.org/10.1016/j.chb.2016.07.060>
- Fidalgo-Blanco, Á., Sein-Echaluce, M. L., & García-Peñalvo, F. J. (2013). MOOC cooperativo. Una integración entre cMOOC y xMOOC. In Á. Fidalgo-Blanco & M. L.

- Sein-Echaluze (Eds.), *Actas del II Congreso Internacional sobre Aprendizaje, Innovación y Competitividad, CINAIC 2013 (Madrid, 6-8 de noviembre de 2013)* (pp. 481–486). Madrid, España: Fundación General de la Universidad Politécnica de Madrid.
- Fidalgo-Blanco, Á., Sein-Echaluze, M. L., & García-Peñalvo, F. J. (2018). Ontological Flip Teaching: a Flip Teaching model based on knowledge management. *Universal Access in the Information Society*, 17(3), 475–489. <https://doi.org/10.1007/s10209-017-0556-6>
- Fidalgo-Blanco, Á., Sein-Echaluze, M. L., & García Peñalvo, F. J. (2017). Inteligencia Colectiva en el aula. Un paradigma cooperativo. In Á. Fidalgo-Blanco, M. L. Sein-Echaluze, & F. J. García Peñalvo (Eds.), *La innovación docente como misión del profesorado. Actas del IV Congreso Internacional sobre Aprendizaje, Innovación y Competitividad. CINAIC 2017 (4-6 de Octubre de 2017, Zaragoza, España)* (pp. 599–603). Zaragoza, España: Servicio de Publicaciones Universidad de Zaragoza.
- Flanagan, B., & Ogata, H. (2018). Learning analytics platform in higher education in Japan. *Knowledge Management and E-Learning*, 10(4), 469–484.
- Fonseca, D., Pifarre, M., Redondo, E., Alitany, A., & Sanchez, A. (2013). Combination of qualitative and quantitative techniques in the analysis of new technologies implementation in education: Using augmented reality in the visualization of architectural projects. *Iberian Conference on Information Systems and Technologies, CISTI*, 1–7.
- Francis, L. P. (2008). Privacy and Confidentiality: the Importance of Context. *Monist*, Vol. 91, pp. 52–67. <https://doi.org/10.2307/27904065>
- García-Holgado, A., & García-Peñalvo, F. J. (2017). Definición de ecosistemas de aprendizaje independientes de plataforma. In M. L. S.-E. Laclea, Á. Fidalgo-Blanco, & F. J. García-Peñalvo (Eds.), *La innovación docente como misión del profesorado. Actas del IV Congreso Internacional sobre Aprendizaje, Innovación y Competitividad. CINAIC 2017 (4-6 de Octubre de 2017, Zaragoza, España)* (pp. 668–673). https://doi.org/10.26754/cinaic.2017.000001_143

- García-Peñalvo, F. J. (2005). Estado actual de los sistemas e-learning. Retrieved July 24, 2018, from Education in the knowledge society (EKS) website: http://campus.usal.es/~teoriaeducacion/rev_numero_06_2/n6_02_art_garcia_peñalvo.htm
- García-Peñalvo, F. J. (2008). Advances in e-learning: Experiences and methodologies. In *Advances in E-Learning: Experiences and Methodologies*. <https://doi.org/10.4018/978-1-59904-756-0>
- García-Peñalvo, F. J. (2013). Education in Knowledge Society. A new PhD Programme approach. In F. J. García-Peñalvo (Ed.), *Proceedings of the First International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'13) (Salamanca, Spain, November 14-15, 2013)* (pp. 575–577). <https://doi.org/10.1145/2536536.2536624>
- García-Peñalvo, F. J. (2014). Formación en la sociedad del conocimiento, un programa de doctorado con una perspectiva interdisciplinar. *Teoría de La Educación. Educación y Cultura En La Sociedad de La Información*, 15(1), 4–9.
- García-Peñalvo, F. J. (2015). Mapa de tendencias en Innovación Educativa. *Education in the Knowledge Society (EKS)*, 16(4), 6. <https://doi.org/10.14201/eks2015164623>
- García-Peñalvo, F. J. (2016a). ¿Son conscientes las universidades de los cambios que se están produciendo en la Educación Superior? *Education in the Knowledge Society (EKS)*, 17(4), 7. <https://doi.org/10.14201/eks2016174713>
- García-Peñalvo, F. J. (2016b, March 7). Presentation of the GRIAL research group and its main research lines and projects on March 2016. Retrieved July 1, 2018, from <https://gredos.usal.es/handle/10366/127737>
- García-Peñalvo, F. J. (2018). Ecosistemas tecnológicos universitarios. In J. Gómez (Ed.), *Análisis de las TIC en las Universidades Españolas* (pp. 164–170). Madrid, España: Crue Universidades Españolas.
- García-Peñalvo, F. J. (2019a). La transformación digital de la docencia. Retrieved June 1, 2019, from <https://bit.ly/2MQUUbe>
- García-Peñalvo, F. J. (2019b). Programa de Doctorado Formación en la Sociedad del

- Conocimiento. Kick-off de la Edición 2019-2020. *Paper Presented at the Seminarios Del Programa de Doctorado En Formación En La Sociedad Del Conocimiento (21 de Octubre de 2019)*. <https://doi.org/10.13140/RG.2.2.10624.79368>
- García-Peñalvo, F. J. (2019c). Revisiones y mapeos sistemáticos de literatura. <https://doi.org/10.5281/zenodo.2586725>
- García-Peñalvo, F. J., Fernández-Hermo, V., Fidalgo-Blanco, Á., & Sein-Echaluce, M. L. (2014). Applied educational innovation MOOC: Learners' experience and valorization of strengths and weaknesses. In F. J. García-Peñalvo (Ed.), *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM 2014) (Salamanca, Spain, October 1-3, 2014)* (pp. 139–145). <https://doi.org/10.1145/2669711.2669892>
- García-Peñalvo, F. J., Fidalgo-Blanco, Á., & Sein-Echaluce, M. L. (2018). An adaptive hybrid MOOC model: Disrupting the MOOC concept in higher education. *Telematics and Informatics*, 35(4), 1018–1030. <https://doi.org/10.1016/j.tele.2017.09.012>
- García-Peñalvo, F. J., Fidalgo-Blanco, Á., Sein-Echaluce, M. L., & Sánchez-Canales, M. (2019). Active Peer-Based Flip Teaching: An Active Methodology Based on RT-CICLO. In M. L. Sein-Echaluce, Á. Fidalgo-Blanco, & F. J. García-Peñalvo (Eds.), *Innovative Trends in Flipped Teaching and Adaptive Learning* (pp. 1–16). <https://doi.org/10.4018/978-1-5225-8142-0.ch001>
- García-Peñalvo, F. J., Hernández-García, Á., Conde, M. Á., Fidalgo-Blanco, Á., Sein-Echaluce, M. L., Alier, M., ... Iglesias-Pradas, S. (2017). Enhancing education for the knowledge society era with learning ecosystems. In F. J. García-Peñalvo & A. García-Holgado (Eds.), *Open Source Solutions for Knowledge Management and Technological Ecosystems* (pp. 1–24). <https://doi.org/10.4018/978-1-5225-0905-9.ch001>
- García-Peñalvo, F. J., Hernández-García, Á., Conde, M. Á., Fidalgo-Blanco, Á., Sein-Echaluce, M. L., Alier, M., ... Iglesias-PRadas, S. (2015). Mirando hacia el futuro: Ecosistemas tecnológicos de aprendizaje basados en servicios Looking into the future: Learning services-based technological ecosystems. In Á. Fidalgo-Blanco, M. L. Sein-Echaluce, & F. J. García-Peñalvo (Eds.), *La Sociedad del Aprendizaje. Actas*

del III Congreso Internacional sobre Aprendizaje, Innovación y Competitividad. CINAIC 2015 (14-16 de Octubre de 2015, Madrid, España) (pp. 553–558). Madrid, Spain: Fundación General de la Universidad Politécnica de Madrid.

García-Peñalvo, F. J., & Seoane-Pardo, A. M. (2015). Una revisión actualizada del concepto de eLearning. Décimo Aniversario. *Education in the Knowledge Society (EKS)*, 16(1), 119–144. <https://doi.org/10.14201/eks2015161119144>

García Aretio, L. (2017). Los MOOC están muy vivos. Respuestas a algunas preguntas. *RIED. Revista Iberoamericana de Educación a Distancia*, 20(1), 9. <https://doi.org/10.5944/ried.20.1.17488>

Gartner. (2019). Gartner 2019 Hype Cycle for Blockchain Business Shows Blockchain Will Have a Transformational Impact across Industries in Five to 10 Years. Retrieved September 14, 2019, from Newsroom website: <https://www.gartner.com/en/newsroom/press-releases/2019-09-12-gartner-2019-hype-cycle-for-blockchain-business-shows>

Gartner Inc. (2016). Hype Cycle Research Methodology. Retrieved September 7, 2019, from Gartner Inc. website: <http://www.gartner.com/technology/research/methodologies/hype-cycle.jsp>

Gervais, A., Karame, G. O., Wüst, K., Glykantzis, V., Ritzdorf, H., & Čapkun, S. (2016). On the security and performance of Proof of Work blockchains. *Proceedings of the ACM Conference on Computer and Communications Security, 24-28-Octo*, 3–16. <https://doi.org/10.1145/2976749.2978341>

Gilda, S., & Mehrotra, M. (2018). Blockchain for Student Data Privacy and Consent. *2018 International Conference on Computer Communication and Informatics, ICCCI 2018*, 1–5. <https://doi.org/10.1109/ICCCI.2018.8441445>

Glance, D. G., Forsey, M., & Riley, M. (2013). The pedagogical foundations of massive open online courses. *First Monday*, 18(5). <https://doi.org/10.5210%2Ffm.v18i5.4350>

Glasserman, L., Mortera, F., & Montoya, M. S. (2013). Caracterizando recursos educativos abiertos (REA) y objetos de aprendizaje (OA) que fomentan un

- aprendizaje activo en los alumnos de primaria. In F. J. Mortera & M. S. Ramírez-Montoya (Eds.), *Conexión de repositorios educativos digitales: Educonector.info* (pp. 26–34). México: Lulú editorial digital.
- Gobierno de España. (2011). Real Decreto 99/2011, de 28 de enero, por el que se regulan las enseñanzas oficiales de doctorado. In *Ministerio de Educación* (Vol. 35, pp. 13909–13926). Gobierno de España.
- Goldberg, I. (2007). Privacy-enhancing technologies for the internet III: Ten years later. In *Digital Privacy: Theory, Technologies, and Practices* (pp. 3–18). Auerbach Publications.
- Goldberg, I., Wagner, D., & Brewer, E. (1997). Privacy-enhancing technologies for the Internet. *Digest of Papers - COMPCON - IEEE Computer Society International Conference*, 103–109.
- Goldstein, P. J., & Katz, R. N. (2005). Academic Analytics : The Uses Of Management Information And Technology In Higher Education. <https://doi.org/10.1080/17439880802097659>
- Gómez-Aguilar, D. A., García-Peñalvo, F. J., & Therón, R. (2014). Analítica visual en e-learning. *Profesional de La Informacion*, 23(3), 236–245. <https://doi.org/10.3145/epi.2014.may.03>
- Gómez-Aguilar, D. A., Hernández-García, Á., García-Peñalvo, F. J., & Therón, R. (2015). Tap into visual analysis of customization of grouping of activities in eLearning. *Computers in Human Behavior*, 47, 60–67. <https://doi.org/10.1016/j.chb.2014.11.001>
- Gong, X., Liu, X., Jing, S., Xiong, G., & Zhou, J. (2019). Parallel-Education-Blockchain Driven Smart Education: Challenges and Issues. *Proceedings 2018 Chinese Automation Congress, CAC 2018*, 2390–2395. <https://doi.org/10.1109/CAC.2018.8623198>
- Gracia, F. A., Gil, P. O., & Osinaga, A. A. (2012). Social learning in formal education through peer collaborative networks and activities: research results and hazards. *X Jornades de Xarxes d'Investigació En Docència Universitària: La Participació i El*

Compromiss de La Comunitat Universitària, 928–938.

- Graham, C. R. (2006). Blended learning systems: Definition, current trends, and future directions. In C. J. Bonk & C. R. Graham (Eds.), *Handbook of blended learning: Global perspectives, local designs* (pp. 3–21). San Francisco, USA: JosseyBass/Pfeiffer.
- Granollers, T., Perdrix, F., & Lorés, J. (2004). Incorporación de usuarios en la evaluación de la usabilidad por recorrido cognitivo. *Interacción'04*.
- Greller, W., & Drachsler, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Journal of Educational Technology & Society*, 15(3), 42–57. <https://doi.org/10.2307/jeductechsoci.15.3.42>
- Gros, B., & García-Peñalvo, F. J. (2016). Future Trends in the Design Strategies and Technological Affordances of E-Learning. *Learning, Design, and Technology*, 1–23. https://doi.org/10.1007/978-3-319-17727-4_67-1
- Grupo GRIAL. (2019). Producción Científica del Grupo GRIAL de 2011 a 2019. Retrieved from (GRIAL-TR-2019-010) Salamanca, España: Grupo GRIAL, Universidad de Salamanca. website: <https://bit.ly/30I9mLh>
- Guàrdia, L., Maina, M., & Sangrà, A. (2013). MOOC design principles: A pedagogical approach from the learner's perspective. *ELearning Papers*, (33).
- Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, 29(7), 1645–1660. <https://doi.org/10.1016/j.future.2013.01.010>
- Haggard, S. (2013). *The maturing of the MOOC: Literature review of massive open online courses and other forms of online distance learning*. 1–123. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/240193/13-1173-maturing-of-the-mooc.pdf
- Hamre, L., & Vidgen, R. (2008). Using social networks and communities of practice to support information systems implementation. *16th European Conference on Information Systems, ECIS 2008*, 2052–2063.
- Han, M., Wu, D., Li, Z., Xie, Y., He, J. S., & Baba, A. (2018). A novel blockchain-based education records verification solution. *SIGITE 2018 - Proceedings of the 19th*

- Annual SIG Conference on Information Technology Education*, 178–183.
<https://doi.org/10.1145/3241815.3241870>
- Harmelen, M. van, & Workman, D. (2012). Analytics for Learning and Teaching. *CETIS Analytics Series*, 1(3), 1–40. Retrieved from <http://publications.cetis.ac.uk/wp-content/uploads/2012/11/Analytics-for-Learning-and-Teaching-Vol1-No3.pdf>
- Hassenzahl, M., & Tractinsky, N. (2006). User experience-a research agenda. *Behaviour & Information Technology*, 25(2), 91–97.
- Herold, B. (2014). InBloom to Shut Down Amid Growing Data-Privacy Concerns. Retrieved June 21, 2018, from Education Week website: http://blogs.edweek.org/edweek/DigitalEducation/2014/04/inbloom_to_shut_down_amid_growing_data_privacy_concerns.html
- Hill, P. (2016). New Release of European LMS Market Report. Retrieved June 1, 2018, from <https://eliterate.us/new-release-european-lms-market-report/>
- Hoel, T., & Chen, W. (2016). Implications of the European Data Protection Regulations for Learning Analytics Design. In *Proceedings of The International Workshop on Learning Analytics and Educational Data Mining (LAEDM 2016) In conjunction with CRIWG/CollabTech 2016*. Retrieved from http://www.hoel.nu/files/LAEDM_Kanazawa_Sep2016_Hoel_Chen_final_w_header.pdf
- Hogan, A. (2019). Review of Ben Williamson (2017). Big Data in Education: the Digital Future of Learning, Policy and Practice. *Postdigital Science and Education*. <https://doi.org/10.1007/s42438-019-00059-6>
- Hurricane Electric. (2019). Hurricane Electric Internet Services. Retrieved November 1, 2018, from <https://he.net/>
- Iansiti, M., & Lakhani, K. R. (2017). The truth about blockchain. <https://doi.org/10.1016/j.annals.2005.11.001>
- IBTAC. (2019). IBTAC19.
- Inrupt Inc. (2019). Solid. Retrieved September 19, 2019, from [inrupt.com](https://solid.inrupt.com/) website: <https://solid.inrupt.com/>

- Intelliboard. (2019). Intelliboard.net. Retrieved June 1, 2018, from <https://intelliboard.net/>
- Isaak, J., & Hanna, M. J. (2018). User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection. *Computer*, 51(8), 56–59. <https://doi.org/10.1109/MC.2018.3191268>
- Islas-Carmona, J. O. (2008). El prosumidor. El actor comunicativo de la sociedad de la ubicuidad. *Palabra Clave*, 11(1), 29–39.
- James Mazoue. (2013). The MOOC Model: Challenging Traditional Education. Retrieved September 17, 2019, from Educause Review website: <https://er.educause.edu/articles/2013/1/the-mooc-model-challenging-traditional-education>
- Jing, M. (2019). BrainCo CEO says his ‘mind-reading’ tech is here to improve concentration, not surveillance. Retrieved June 7, 2019, from South China Morning Post website: <https://www.scmp.com/tech/innovation/article/3008439/brainco-ceo-says-his-mind-reading-tech-here-improve-concentration>
- Joksimović, S., Kovanović, V., & Dawson, S. (2019). The Journey of Learning Analytics. In *HERDSA Review of Higher Education*, Vol. 6 (p. 89). Retrieved from <http://www.herdsa.org.au/herdsa-review-higher-education-vol-6/37-63>
- Joshi, A. P., Han, M., & Wang, Y. (2018). A survey on security and privacy issues of blockchain technology. *Mathematical Foundations of Computing*, 1(2), 121–147.
- Just Keynote. (2019). Retrieved June 1, 2019, from JUSTKEYNOTE website: <https://justkeynote.com/>
- Kharif, O., & Marsh, A. (2019). Binance CEO Spurs Outcry by Suggesting Blockchain Rollback. Retrieved from Bloomberg website: <https://www.bloomberg.com/news/articles/2019-05-08/crypto-savior-spurs-outcry-by-suggesting-blockchain-rollback>
- Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering -- A systematic literature review. *Information and Software Technology*, 51(1), 7–15.

<https://doi.org/10.1016/j.infsof.2008.09.009>

Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. *Technical Report, Ver. 2.3 EBSE Technical Report. EBSE, EBSE-2007-(School of Computer Science and Mathematics)*, 65. Retrieved from https://www.elsevier.com/__data/promis_misc/525444systematicreviewsguide.pdf

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. *ACM International Conference Proceeding Series*, 170–179. <https://doi.org/10.1145/2460296.2460330>

Kraemer, K. L., Dedrick, J., & Sharma, P. (2009). One laptop per child: vision vs. reality. *Communications of the ACM*, 52(6), 66–73.

L'Heureux, A., Grolinger, K., Elyamany, H. F., & Capretz, M. A. M. (2017). Machine Learning with Big Data: Challenges and Approaches. *IEEE Access*, 5, 7776–7797. <https://doi.org/10.1109/ACCESS.2017.2696365>

Landwehr, C. E. (2018). Computer Security. *International Journal of Information Security*, 10683(1), 3–13. <https://doi.org/10.1007/978-3-319-72817-9>

Lang, C., Macfadyen, L. P., Slade, S., Prinsloo, P., & Sclater, N. (2018). The complexities of developing a personal code of ethics for learning analytics practitioners implications for institutions and the field. *ACM International Conference Proceeding Series*, 436–440. <https://doi.org/10.1145/3170358.3170396>

Lang, C., Siemens, G., Wise, A., & Gasevic, D. (2017). Handbook of Learning Analytics. In *Handbook of Learning Analytics*. <https://doi.org/10.18608/hla17>

Laveti, R. N., Kuppili, S., Ch, J., Pal, S. N., & Babu, N. S. C. (2017). Implementation of learning analytics framework for MOOCs using state-of-the-art in-memory computing. *E-Learning & E-Learning Technologies (ELELTECH), 2017 5th National Conference On*, 1–6.

Learning Analytics. (2010). Retrieved April 2, 2019, from Wikipedia website:

- https://en.wikipedia.org/wiki/Learning_analytics
- Learning Analytics Google Groups. (2010). Retrieved April 27, 2019, from <https://groups.google.com/forum/#!forum/learninganalytics>
- Lee, V. R. (2014). What's happening in the "Quantified Self" movement? *ICLS 2014 Proceedings*, 1032.
- Leenes, R., & Kosta, E. (2015). Taming the cookie monster with Dutch law - A tale of regulatory failure. *Computer Law and Security Review*, 31(3), 317–335. <https://doi.org/10.1016/j.clsr.2015.01.004>
- Li, Y. (2019). MOOCs in Higher Education: Opportunities and Challenges. *5th International Conference on Humanities and Social Science Research (ICHSSR 2019)*. <https://doi.org/https://doi.org/10.2991/ichssr-19.2019.10>
- Liberatori, M., Otero, F., Bonaadero, J. C., & Castiñeira, J. (2007). AES-128 cipher. high speed, low cost FPGA implementation. *Proceedings - 2007 3rd Southern Conference on Programmable Logic, SPL'07*, 195–198. <https://doi.org/10.1109/SPL.2007.371748>
- Litan, A., & Leow, A. (2019). Hype Cycle for Blockchain Technologies. Retrieved September 7, 2019, from Gartner Inc. website: <https://www.gartner.com/en/documents/3947355/hype-cycle-for-blockchain-technologies-2019>
- Liyanagunawardena, T., Williams, S., & Adams, A. (2013). The impact and reach of MOOCs: a developing countries' perspective. *ELearning Papers*, 38–46.
- Llorca, J., Zapata, H., Redondo, E., Alba, J., & Fonseca, D. (2018). Bipolar laddering assessments applied to urban acoustics education. *Advances in Intelligent Systems and Computing*, 747, 287–297. https://doi.org/10.1007/978-3-319-77700-9_29
- Llorens, F., Molina, R., Compañ, P., & Satorre, R. (2014). Technological ecosystem for open education. In R. Neves-Silva, G. A. Tsihrintzis, V. Uskov, R. J. Howlett, & L. C. Jain (Eds.), *Smart Digital Futures 2014* (Vol. 262, pp. 706–715). Amsterdam, The Netherlands: IOS Press.
- Long, P., & Siemens, G. (2011). Penetrating the Fog: Analytics in Learning and Education.

- Retrieved July 24, 2018, from EDUCAUSE Review website: <https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education>
- Long, P., Siemens, G., Gráinne, C., & Gašević, D. (2011). LAK '11 : proceedings of the 1st International Conference on Learning Analytics and Knowledge, February 27 - March 1, 2011, Banff, Alberta, Canada. In *1st International Conference on Learning Analytics and Knowledge*. Retrieved from <https://dl.acm.org/citation.cfm?id=2090116>
- Luigi, A., Antonio, I., & Giacomo, M. (2010). The internet of things: a survey. *Computer Networks*, 54(15), 2787–2805. <https://doi.org/10.1016/j.comnet.2010.05.01>
- Lupton, D., & Williamson, B. (2017). The datafied child: The dataveillance of children and implications for their rights. *New Media and Society*, 19(5), 780–794. <https://doi.org/10.1177/1461444816686328>
- Maldonado, J., & Hilliger, I. (2018). *The LALA Project: Building Capacity to Use Learning Analytics to Improve Higher Education in Latin America*. Retrieved from <https://www.researchgate.net/publication/325127346>
- Mañas, J. L. P. (2005). El derecho fundamental a la protección de datos personales. *Protección de Datos de Carácter Personal En Iberoamérica:(II Encuentro Iberoamericano de Protección de Datos, La Antigua-Guatemala, 2-6 de Junio de 2003)*, 19–36.
- Mandinach, E. B., & Gummer, E. S. (2013). A Systemic View of Implementing Data Literacy in Educator Preparation. *Educational Researcher*, 42(1), 30–37. <https://doi.org/10.3102/0013189X12459803>
- Manel, J., & Sanz, A. (2016). *Blockchain per l'educació*. Retrieved from https://ddd.uab.cat/pub/tfg/2017/tfg_71006/Article_JoanManelArcas.pdf
- Marcucci, P. N., & Johnstone, D. B. (2007). Tuition fee policies in a comparative perspective: Theoretical and political rationales. *Journal of Higher Education Policy and Management*, 29(1), 25–40. <https://doi.org/10.1080/13600800600980015>
- Martinez-Maldonado, R., Hernandez-Leo, D., Pardo, A., Suthers, D., Kitto, K., Charleer,

- S., ... Ogata, H. (2016). Cross-LAK: learning analytics across physical and digital spaces. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, 486–487. Retrieved from <https://dl.acm.org/citation.cfm?id=2883855>
- Mashey, J. R. (1998). Big Data and the Next Wave of InfraStress Problems, Solutions, Opportunities. Retrieved September 7, 2019, from USENIX website: <https://www.usenix.org/conference/1999-usenix-annual-technical-conference/big-data-and-next-wave-infrastress-problems>
- Matcha, W., Ahmad Uzir, N., Gasevic, D., & Pardo, A. (2019). A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. *IEEE Transactions on Learning Technologies*, 1–1. <https://doi.org/10.1109/tlt.2019.2916802>
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.
- Mazoue, J. G., & James Mazoue. (2013). The MOOC Model : Challenging Traditional Education. *EDUCAUSE Review*, 1–9. Retrieved from <https://er.educause.edu/articles/2013/1/the-mooc-model-challenging-traditional-education>
- Mense, A., & Flatscher, M. (2018). Security Vulnerabilities in Ethereum Smart Contracts. *Proceedings of the 20th International Conference on Information Integration and Web-Based Applications & Services*, 375–380.
- Meyer, J., Gurrin, C., Simske, S., Hermens, H. J., & Siek, K. (2014). Beyond quantified self: Data for wellbeing. *Conference on Human Factors in Computing Systems - Proceedings*, 95–98. <https://doi.org/10.1145/2559206.2560469>
- Millard, C. (2018). Blockchain and law: Incompatible codes? *Computer Law & Security Review*, 34(4), 843–846.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Molins, A. (2019). Catalunya pone en marcha su propio sistema de identidad digital

- basado en blockchain. Retrieved September 7, 2019, from La Vanguardia website: <https://www.lavanguardia.com/tecnologia/actualidad/20190907/47200867814/identificat-identidad-digital-catalunya-blockchain.html>
- Moodle HQ, Pataleta, J., & Monllaó, D. (2018). Moodle Plugin Data Privacy. Retrieved October 1, 2018, from Moodle website: https://moodle.org/plugins/tool_dataprivacy
- Morueta, R. T., Gómez, Á. H., & Gómez, J. I. A. (2011). Aprendizaje cooperativo on-line a través de foros en un contexto universitario: Un análisis del discurso y de las redes. *Estudios Sobre Educacion*, (20), 49–71.
- Moubarak, J., Filiol, E., & Chamoun, M. (2018). On blockchain security and relevant attacks. *2018 IEEE Middle East and North Africa Communications Conference, MENACOMM 2018*, 1–6. <https://doi.org/10.1109/MENACOMM.2018.8371010>
- Muñoz, C., García-Peñalvo, F. J., Morales, E. M., Conde, M. Á., & Seoane, A. M. (2012). Improving learning object quality: Moodle HEODAR implementation. *International Journal of Distance Education Technologies*, 10(4), 1–16. <https://doi.org/10.4018/jdet.2012100101>
- Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. 1–9. Retrieved from <https://www.bitcoincash.org/bitcoin.pdf>
- Narayanan, A. (2015). “Private blockchain” is just a confusing name for a shared database. Retrieved from Freedom to tinker website: <https://freedom-to-tinker.com/2015/09/18/private-blockchain-is-just-a-confusing-name-for-a-shared-database/>
- Neiva, F. W., David, J. M. N., Braga, R., & Campos, F. (2016). Towards pragmatic interoperability to support collaboration: A systematic review and mapping of the literature. *Information and Software Technology*, 72, 137–150. <https://doi.org/10.1016/j.infsof.2015.12.013>
- Next Generation. (2010). Learning Analytics. Retrieved April 27, 2019, from Next Gen Learning Challenges website: <http://web.archive.org/web/20100711081424/http://www.nextgenlearning.com/>

the-challenges/learning-analytics

NLTK Project. (2019). Natural Language Toolkit — NLTK 3.4.3 documentation. Retrieved April 27, 2019, from <https://www.nltk.org/>

O'Reilly, T. (2007). What is web 2.0?: Design patterns and business models for the next generation of software. *Communications & Strategies*, 1(65), 17–37.

Ocheja, P., Flanagan, B., & Ogata, H. (2018). Connecting decentralized learning records: a blockchain based learning analytics platform. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, 265–269. <https://doi.org/10.1145/3170358.3170365>

Ochoa, X., Lang, A. C., & Siemens, G. (2017). Multimodal learning analytics. *The Handbook of Learning Analytics*, 1, 129–141. Retrieved from https://www.researchgate.net/profile/Dragan_Gasevic/publication/324687610_Handbook_of_Learning_Analytics/links/5add21e1aca272fdaf86c95c/Handbook-of-Learning-Analytics.pdf#page=129

Ochoa, X., Weibel, N., Worsley, M., & Oviatt, S. (2016). Multimodal learning analytics data challenges. *6th International Conference on Learning Analytics and Knowledge, LAK 2016*, 498–499. Retrieved from <https://sci-hub.se/https://nyu-staging.pure.elsevier.com/en/publications/multimodal-learning-analytics-data-challenges>

Ochoa, X., & Worsley, M. (2016). Augmenting Learning Analytics with Multimodal Sensory Data. *Journal of Learning Analytics*, 3(2), 213–219. <https://doi.org/10.18608/jla.2016.32.10>

OECD. (2012). *Education at a Glance 2012*. <https://doi.org/10.1787/eag-2012-en>

Ojanen, T. (2016). Making the essence of fundamental rights real: The court of justice of the European Union clarifies the structure of fundamental rights under the charter. *European Constitutional Law Review*, 12(2), 318–329. <https://doi.org/10.1017/S1574019616000225>

Pagallo, U., Bassi, E., Crepaldia, M., & Durante, M. (2018). Chronicle of a clash foretold: Blockchains and the GDPR's right to erasure. *Frontiers in Artificial Intelligence and*

- Applications*, 313, 81–90. <https://doi.org/10.3233/978-1-61499-935-5-81>
- Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, 45(3), 438–450. <https://doi.org/10.1111/bjet.12152>
- Parker, G., Van Alstyne, M., & Choudary, S. (2016). Platform revolution: How networked markets are transforming the economy and how to make them work for you. In *W.W. Norton & Company*. WW Norton & Company.
- Patents, G. (2018). Decentralized credentials verification network. *US Patent App. 15 ...* Retrieved from <https://patents.google.com/patent/US20180082256A1/en>
- Pauner-Chulvi, C., & Viguri Cordero, J. A. (2018). *The Adaptation Of The GDPR In Spain: The New Data Protection Act.* (June). Retrieved from <http://www.congreso.es/portal/page/portal/Congreso/>
- Pedreño, A., Moreno, L., Ramón, A., & Pernías, P. (2013). UniMOOC: trabajo colaborativo e innovación educativa. *Campus Virtuales*, 2(1).
- Peñalvo, F. J. G., Conde, M. J. R., Pardo, A. M. S., González, M. Á. C., Zangrando, V., & Holgado, A. G. (2012). GRIAL (GRupo de investigación en InterAcción y eLearning), USAL. *IE Comunicaciones: Revista Iberoamericana de Informática Educativa*, (15), 85–94.
- Penuel, W. R., Korbak, C., & Hoadley, C. (2006). Investigating the potential of using social network analysis in educational evaluation. *American Journal of Evaluation*, 27(4), 437–451. <https://doi.org/10.1177/1098214006294307>
- Petticrew, M., & Roberts, H. (2008). Systematic Reviews in the Social Sciences: A Practical Guide. In *Systematic Reviews in the Social Sciences: A Practical Guide*. <https://doi.org/10.1002/9780470754887>
- Pifarré, M., & Tomico, O. (2007). Bipolar laddering (BLA): A participatory subjective exploration method on user experience. *Proceedings of the 2007 Conference on Designing for User Experiences, DUX'07*, 2. <https://doi.org/10.1145/1389908.1389911>
- Popa, R. A., Redfield, C. M. S., Zeldovich, N., & Balakrishnan, H. (2012). CryptDB:

- Processing queries on an encrypted database. *Communications of the ACM*, 55(9), 103–111. <https://doi.org/10.1145/2330667.2330691>
- Poveda, L. A. (2018). Alguns aspectes sobre blockchains i smart contracts en educació superior. *Revista d'Innovació Docent Universitària*, 0(10), 65–76. <https://doi.org/10.1344/RIDU2018.10.7>
- Preibusch, S., Krol, K., & Beresford, A. R. (2013). The privacy economics of voluntary over-disclosure in web forms. In *The Economics of Information Security and Privacy* (pp. 183–209). https://doi.org/10.1007/978-3-642-39498-0_9
- Preneel, B. (2010). The first 30 years of cryptographic hash functions and the NIST SHA-3 competition. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5985 LNCS, 1–14. https://doi.org/10.1007/978-3-642-11925-5_1
- Pressman, R. S., & Maxim, B. R. (2015). *Software Engineering: A practitioner's approach* (8th ed.). New York, NY, USA: McGraw-Hill Education.
- Prinsloo, T., & Van Deventer, J. P. (2017). Using the Gartner Hype Cycle to evaluate the adoption of emerging technology trends in higher education – 2013 to 2016. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 10676 LNCS* (pp. 49–57). https://doi.org/10.1007/978-3-319-71084-6_7
- Ramírez-Montoya, M. S. (2015). Acceso abierto y su repercusión en la Sociedad del Conocimiento: Reflexiones de casos prácticos en Latinoamérica. *Education in the Knowledge Society (EKS)*, 16(1), 103–118. <https://doi.org/10.14201/eks2015161103118>
- Ramírez-Montoya, M. S., & García-Peñalvo, F. J. (2015). Movimiento Educativo Abierto. *Virtualis*, 6(12), 1–13. Retrieved from <http://aplicaciones.ccm.itesm.mx/virtualis/index.php/virtualis/article/view/125>
- Ramírez Gómez, M. (2018). Guia de protecció de dades en centres educatius. Retrieved June 1, 2018, from <http://compromesosambleducacio.diba.cat/blogs/2018/02/20/guia-de-proteccio->

de-dades-centres-educatius

- Razaghpanah, A., Nithyanand, R., Vallina-Rodriguez, N., Sundaresan, S., Allman, M., Kreibich, C., & Gill, P. (2018). *Apps, Trackers, Privacy, and Regulators: A Global Study of the Mobile Tracking Ecosystem*. <https://doi.org/10.14722/ndss.2018.23353>
- Razaghpanah, A., Vallina-Rodriguez, N., Sundaresan, S., Kreibich, C., Gill, P., Allman, M., & Paxson, V. (2015). Haystack: In situ mobile traffic analysis in user space. *ArXiv Preprint ArXiv:1510.01419*, 1–13.
- Read, J. C. (2006). A study of the usability of handwriting recognition for text entry by children. *Interacting with Computers*, 19(1), 57–69.
- Robinson, M. (2017). AltSchool, funded by tech execs, is closing schools, losing students - Business Insider. Retrieved April 7, 2019, from Business Insider website: <https://www.businessinsider.com/altschool-why-parents-leaving-2017-11?IR=T>
- Rodríguez, E. (2019). *El uso de Blockchain para resguardar analíticas de aprendizaje* (Universitat Politècnica de Catalunya). Retrieved from <https://upcommons.upc.edu/bitstream/handle/2117/133059/138146.pdf>
- Rodriguez, O. (2012). MOOCs and the AI-Stanford like Courses: two successful and distinct course formats for massive open online courses. *European Journal of Open, Distance, and E-Learning*, 1–13. Retrieved from <http://www.eurodl.org/materials/contrib/2012/Rodriguez.htm>
- Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, 33(1), 135–146. <https://doi.org/10.1016/j.eswa.2006.04.005>
- Romero, C., & Ventura, S. (2010). Educational Data Mining: A Review of the State of the Art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(6), 601–618. <https://doi.org/10.1109/TSMCC.2010.2053532>
- Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 12–27. <https://doi.org/10.1002/widm.1075>
- Rosenbaum, J. I. (1997). Privacy on the internet: whose information is it anyway.

Jurimetrics, 38(4), 565–573. Retrieved from <http://www.jstor.org/stable/29762571>

Sakurai, Y. (2014). The value improvement in education service by grasping the value acceptance state with ICT utilized education environment. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8522 LNCS(PART 2), 90–98. https://doi.org/10.1007/978-3-319-07863-2_10

Salazar Rodríguez, A., Rodríguez Gómez, J., & Campos Madrigal, S. (2012). Recursos educativos abiertos y estrategias de búsqueda e implementación en un ambiente de aprendizaje universitario. *EduTec: Revista Electrónica de Tecnología Educativa*, 0(41), 3. <https://doi.org/10.21556/edutec.2012.41.350>

Sánchez Bravo, A. (1994). La regulación de los datos sensibles en la LORTAD. *Informática y Derecho*, 6-7, 117-132.

Santiago, D. (2016). Privacy vs. Data Protection vs. Information Security. Retrieved June 7, 2019, from Software and Services Engineering by STRAST website: <https://blogs.upm.es/sse/2016/11/01/privacy-vs-data-protection-vs-information-security/>

Schneier, B., & Hardie, T. (2014). Pervasive Attack: A Threat Model and Problem Statement. Retrieved June 1, 2018, from IETF website: <https://datatracker.ietf.org/doc/html/draft-barnes-pervasive-problem>

Sclater, N. (2008). Web 2.0, personal learning environments, and the future of learning management systems. *Research Bulletin*, 13(13), 1–13.

Sclater, N., & Biley, P. (2015). Code of practice for learning analytics | Jisc. Retrieved November 22, 2018, from JISC website: <https://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics>

Sein-Echaluce, M. L., Fidalgo-Blanco, A., & García-Peñalvo, F. J. (2019). *Innovative trends in flipped teaching and adaptive learning*. Hershey, PA, USA: IGI Global.

Sein-Echaluce, M. L., Fidalgo-Blanco, Á., & García-Peñalvo, F. J. (2015). Metodología de enseñanza inversa apoyada en b-learning y gestión del conocimiento. In M. L. Sein-

- Echaluze, Á. Fidalgo-Blanco, & F. J. García-Peñalvo (Eds.), *La Sociedad del Aprendizaje. Actas del III Congreso Internacional sobre Aprendizaje, Innovación y Competitividad. CINAIC 2015 (14-16 de Octubre de 2015, Madrid, España)* (pp. 464–468). Retrieved from <https://repositorio.grial.eu/handle/grial/480>
- Seničar, V., Jerman-Blažič, B., & Klobučar, T. (2003). Privacy-enhancing technologies—approaches and development. *Computer Standards & Interfaces*, 25(2), 147–158.
- Shah, D. (2019). Year of MOOC-based Degrees: A Review of MOOC Stats and Trends in 2018. Retrieved September 17, 2019, from Class Central website: <https://www.classcentral.com/report/moocs-stats-and-trends-2018/>
- Shea, P., Hayes, S., Uzuner Smith, S., Vickers, J., Bidjerano, T., Gozza-Cohen, M., ... Tseng, C. H. (2013). Online learner self-regulation: Learning presence viewed through quantitative content- and social network analysis. *International Review of Research in Open and Distance Learning*, 14(3), 427–461.
- Siemens, G. (2005). Connectivism : A Learning Theory for the Digital Age. *International Journal of Instructional Technology and Distance Learning*, 2(1), 3–10.
- Siemens, G. (2007). Connectivism: Creating a Learning Ecology in Distributed Environments. In T. Hug (Ed.), *Didactics of Microlearning Concepts Discourses and Examples* (pp. 53–68). Münster, Germany: Waxmann.
- Siemens, G. (2010a). 1st International Conference on Learning Analytics and Knowledge 2011 | Connecting the technical, pedagogical, and social dimensions of learning analytics. Retrieved September 18, 2019, from Call for Participation for LAK '11 website: <https://tekri.athabascau.ca/analytics/>
- Siemens, G. (2010b). Welcome to the open online course on Learning & Knowledge Analytics. Retrieved April 27, 2019, from Learning and Knowledge Analytics | Analyzing what can be connected website: <https://www.learninganalytics.net/page/4/>
- Siemens, G. (2010c). What are Learning Analytics? Retrieved April 27, 2019, from Elearnspace website: <http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/>

- Siemens, G. (2012). Learning analytics: Envisioning a research discipline and a domain of practice. *LAK '12 Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 4–8. <https://doi.org/10.1145/2330601.2330605>
- Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioral Scientist*, 57(10), 1380–1400. <https://doi.org/10.1177/0002764213498851>
- Siemens, G., & Downes, S. (2008). Connectivism and Connective Knowledge Online Course. Retrieved September 17, 2019, from <https://web.archive.org/web/20080617133303/http://www.elearnspace.org/connectivism.html>
- Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Shum, S. B., Ferguson, R., ... Verbert, K. (2011). Open learning analytics: An integrated & Modularized platform. Proposal to design, implement and evaluate an open platform to integrate heterogeneous learning analytics techniques. Retrieved April 1, 2013, from SOLAR website: <http://solaresearch.org/OpenLearningAnalytics.pdf>
- Singer, N. (2014). InBloom Student Data Repository to Close. Retrieved July 24, 2018, from The New York Times Bit Blog website: http://bits.blogs.nytimes.com/2014/04/21/inbloom-student-data-repository-to-close/?_r=0
- Sobti, R., & Geetha, G. (2012). Cryptographic Hash functions - a review. *International Journal of Computer Science Issues*, 9(2), 461.
- SoLAR. (2019). About SoLAR. Retrieved April 27, 2019, from Society for Learning Analytics Research (SoLAR) website: <https://solaresearch.org/about/>
- Soni, A., & Maheshwari, S. (2018). A Survey of Attacks on the Bitcoin System. *2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science, SCEECS 2018*, 1–5. <https://doi.org/10.1109/SCEECS.2018.8546925>
- Sonwalkar, N. (Nish). (2013). The First Adaptive MOOC: A Case Study on Pedagogy Framework and Scalable Cloud Architecture—Part I. *MOOCs FORUM*, 1(P), 22–29. <https://doi.org/10.1089/mooc.2013.0007>

- Soomro, A. B., Salleh, N., Mendes, E., Grundy, J., Burch, G., & Nordin, A. (2016). The effect of software engineers' personality traits on team climate and performance: A Systematic Literature Review. *Information and Software Technology, 73*, 52–65. <https://doi.org/10.1016/j.infsof.2016.01.006>
- Srivastava, A., Bhattacharya, P., Singh, A., Mathur, A., Prakash, O., & Pradhan, R. (2019). A Distributed Credit Transfer Educational Framework based on Blockchain. *Proceedings - 2018 2nd International Conference on Advances in Computing, Control and Communication Technology, IAC3T 2018*, 54–59. <https://doi.org/10.1109/IAC3T.2018.8674023>
- Stevens, W. R. (1994). *TCP/IP illustrated vol. I: the protocols*. Pearson Education India.
- Stokes, P. (2013). The Particle Accelerator of Learning. Retrieved September 17, 2019, from INSIDE Higher Ed website: <https://www.insidehighered.com/views/2013/02/22/look-inside-edxs-learning-laboratory-essay>
- Sun, H., Wang, X., & Wang, X. (2018). Application of Blockchain Technology in Online Education. *International Journal of Emerging Technologies in Learning (IJET)*, 13(10), 252. <https://doi.org/10.3991/ijet.v13i10.9455>
- Swan, M. (2015). Blockchain: Blueprint for a new economy. In *Climate Change 2013 - The Physical Science Basis*. <https://doi.org/10.1017/CBO9781107415324.004>
- Szabo, N. (1997). Formalizing and securing relationships on public networks. *First Monday*, 2(9). <https://doi.org/10.5210/fm.v2i9.548>
- Szabo, N. (2008). Bit gold. Retrieved June 1, 2018, from Unenumerated website: <https://unenumerated.blogspot.com/2005/12/bit-gold.html>
- Tabaa, Y., & Medouri, A. (2013). LASyM: A Learning Analytics System for MOOCs. *International Journal of Advanced Computer Science and Applications*, 4(5). <https://doi.org/10.14569/ijacsa.2013.040516>
- Taneja, S., & Goel, A. (2014). MOOC Providers and their Strategies. *International Journal of Computer Science and Mobile Computing*, 35(5), 222–228.
- Tapscott, D., & Tapscott, A. (2016). *Blockchain revolution: how the technology behind*

bitcoin is changing money, business, and the world. Penguin.

Telecommunications, C. R. of 17 J. 1995 on the lawful interception of. (n.d.). Lawful Interception - 31996G1104. *Official Journal C 329*, 04/11/1996 P. 0001 - 0006;

Téllez, S. A., & Moodle HQ. (2018). Moodle Plugin Policies. Retrieved October 1, 2018, from Moodle website: https://moodle.org/plugins/tool_policy

Tirado, R., Aguaded, I., & Hernando, A. (2011). Collaborative Learning Processes in an Asynchronous Environment: An Analysis through Discourse and Social Networks. *Online Submission*, 2(1), 115–146.

Tsai, Y.-S., Moreno-Marcos, P. M., Tammets, K., Kollom, K., & Gašević, D. (2018). SHEILA policy framework: informing institutional strategies and policy processes of learning analytics. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge - LAK '18*, 320–329. <https://doi.org/10.1145/3170358.3170367>

Tschofenig, H., & Baccelli, E. (2019). Cyberphysical Security for the Masses: A Survey of the Internet Protocol Suite for Internet of Things Security. *IEEE Security Privacy*, 17(5), 47–57. <https://doi.org/10.1109/MSEC.2019.2923973>

Turcu, C., Turcu, C., & Chiuchișan, I. (2018). Blockchain and its Potential in Education. *International Conference on Virtual Learning - ICVL, Alba Iulia*.

Turkanović, M., Hölbl, M., Košič, K., Heričko, M., & Kamišalić, A. (2018). EduCTX: A blockchain-based higher education credit platform. *IEEE Access*, 6, 5112–5127. <https://doi.org/10.1109/ACCESS.2018.2789929>

UNESCO. (2012). 2012 Paris Oer Declaration. *World Open Educational Resources (OER) Congress*, (June), 2011–2013. Retrieved from http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/Events/English_Paris_OER_Declaration.pdf

Vallina-Rodriguez, N. (2017). *Illuminating the Third Party Mobile Ecosystem with the Lumen Privacy Monitor*. Retrieved from <http://eprints.networks.imdea.org/1522/>

Villagrasa, S., Fonseca, D., & Durán, J. (2014). Teaching case: Applying gamification techniques and virtual reality for learning building engineering 3D arts. *ACM*

- International Conference Proceeding Series*, 171–177.
<https://doi.org/10.1145/2669711.2669896>
- Villagrasa, S., Fonseca, D., Redondo, E., & Duran, J. (2018). Teaching case of gamification and visual technologies for education. *Ophthalmology: Breakthroughs in Research and Practice*, 16(4), 205–226. <https://doi.org/10.4018/978-1-5225-5198-0.ch012>
- Weippl, E. R., & Min Tjoa, A. (2005). Privacy in e-learning: anonymity, pseudonyms and authenticated usage. *Interactive Technology and Smart Education*, 2(4), 247–256. <https://doi.org/10.1108/17415650580000048>
- Weiss, M. A., & Archick, K. (2016). U.S.-EU data privacy: From safe harbor to privacy shield. *The European Union: Challenges and Prospects*, pp. 113–135. Congressional Research Service.
- West, S. M. (2019). Data Capitalism: Redefining the Logics of Surveillance and Privacy. *Business and Society*, 58(1), 20–41. <https://doi.org/10.1177/0007650317718185>
- Wetherall, D., Choffnes, D., Greenstein, B., Han, S., Hornyack, P., Jung, J., ... Wang, X. (2011). Privacy revelations for web and mobile apps. *Proceedings of the 13th USENIX Conference on Hot Topics in Operating Systems*, 21–21. Retrieved from <http://portal.acm.org/citation.cfm?id=1991596.1991625>
- Wexler, S., Dublin, L., Grey, N., Jagannathan, S., Karrer, T., Martinez, M., ... Barneveld, A. v. (2007). LEARNING MANAGEMENT SYSTEMS. The good, the bad, the ugly,... and the truth. In *Guild Research 360 Degree Report*. Santa Rosa, California, USA.
- Wiley, D. A. (2002). Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. *The Instructional Use of Learning Objects: Agency for Instructional Technology*.
- Wilkinson, S., & Lowry, J. (2014). MetaDisk: A Blockchain-Based Decentralized File Storage Application. *Tech. Rep.*, 1–11. Retrieved from <http://metadisk.org/metadisk.pdf>
- Williamson, B. (2016a). Digital education governance: data visualization, predictive analytics, and 'real-time' policy instruments. *Journal of Education Policy*, 31(2), 123–141. <https://doi.org/10.1080/02680939.2015.1035758>

- Williamson, B. (2016b). Digital methodologies of education governance: Pearson plc and the remediation of methods. *European Educational Research Journal*, 15(1), 34–53. <https://doi.org/10.1177/1474904115612485>
- Williamson, B. (2017a). *Big data in education: The digital future of learning, policy and practice* (J. Clark, Ed.). London, UK: SAGE Publications Ltd.
- Williamson, B. (2017b). Decoding ClassDojo: psycho-policy, social-emotional learning and persuasive educational technologies. *Learning, Media and Technology*, 42(4), 440–453. <https://doi.org/10.1080/17439884.2017.1278020>
- Wust, K., & Gervais, A. (2018). Do you need a blockchain? *Proceedings - 2018 Crypto Valley Conference on Blockchain Technology, CVCBT 2018*, 45–54. <https://doi.org/10.1109/CVCBT.2018.00011>
- XNET. (2019). No Firméis la Autorización para el Uso de Google Suite en las Escuelas. Retrieved September 19, 2019, from <https://xnet-x.net/no-autorizar-google-suite-escuelas/>
- Xu, Y., Zhao, S., Kong, L., Zheng, Y., Zhang, S., & Li, Q. (2017). ECBC: A high performance educational certificate blockchain with efficient query. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 10580 LNCS* (pp. 288–304). https://doi.org/10.1007/978-3-319-67729-3_17
- Yang, D., Sinha, T., Adamson, D., & Rose, C. (2013). “Turn on, Tune in, Drop out”: Anticipating student dropouts in Massive Open Online Courses. *Proceedings of the NIPS Workshop on Data Driven Education*, 11, 1–8.
- You, T. (2019). Chinese schools use facial-recognition gates to monitor pupils. Retrieved June 7, 2019, from Daily Mail Online website: <https://www.dailymail.co.uk/news/article-7153981/Chinese-schools-use-facial-recognition-gates-monitor-pupils.html>
- Young, S. (2018). *From Disruption to Innovation*. Retrieved from <https://vo.hse.ru/data/2018/11/19/1141782876/Young.pdf>
- Yu, C.-H., Wu, J., & Liu, A.-C. (2019). Predicting Learning Outcomes with MOOC

- Clickstreams. *Education Sciences*, 9(2), 104.
<https://doi.org/10.3390/educsci9020104>
- Yuan, L., & Powell, S. (2013). MOOCs and disruptive innovation: Implications for higher education. *ELearning Papers*, (33), 1–8. Retrieved from <http://www.openeducationeuropa.eu/en/article/MOOCs-and-disruptive-innovation:-Implications-for-higher-education>
- Zeide, E. (2017). The Structural Consequences of Big Data-Driven Education. *Big Data*, 5(2), 164–172. <https://doi.org/10.1089/big.2016.0061>
- Zheng, Z., Xie, S., Dai, H., Chen, X., & Wang, H. (2017). An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends. *Proceedings - 2017 IEEE 6th International Congress on Big Data, BigData Congress 2017*, 557–564. <https://doi.org/10.1109/BigDataCongress.2017.85>
- Zimmerman, R. K. (2001). The way the “ Cookies ” crumble : Internet privacy Twenty-first Century. *Public Policy*, 43(1997), 439–464. Retrieved from <https://www.merchantgould.com/portalresource/The-Way-the-Cookies-Crumble-Rachel-Zimmerman-Scobie.pdf>
- Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30(1), 75–89. <https://doi.org/10.1057/jit.2015.5>