

Uso de arquitecturas software para recolectar información de interacción en entornos eLearning

Juan Cruz-Benito¹; Francisco José García-Peñalvo¹; Roberto Therón¹; Cristina Maderuelo²; Jonás Samuel Pérez-Blanco²; Hinojal Zazo²; Ana Martín-Suárez²

¹Grupo de Investigación GRIAL
Departamento de Informática y Automática
Universidad de Salamanca
Salamanca, España
{juancb, fgarcia, theron}@usal.es

²Departamento de Farmacia y Tecnología Farmacéutica
Universidad de Salamanca
Salamanca, España
{cmaderuelo, jspez, hinojal, amasu}@usal.es

Resumen—Este trabajo presenta una arquitectura software basada en servicios y desplegada en un entorno *cloud* que recoge, analiza y presenta información extraída de un ecosistema de aprendizaje cerrado como pueden ser los Mundos Virtuales. Esta arquitectura es capaz de recopilar la interacción de los usuarios dentro de la plataforma digital, organizar los datos de dicha interacción y realizar mediciones, estimaciones y análisis básico sobre los datos con el objetivo de dar información a los responsables de la actividad docente acerca de los indicadores de uso del entorno virtual y el grado de resolución de los alumnos de determinados objetivos perseguidos por los docentes que plantean estos escenarios y sistemas de aprendizaje. Para probar esta idea, se presenta la aplicación de la capa de análisis de esta arquitectura sobre un caso real, de modo que se puedan observar cómo podría ayudar una arquitectura basada en capas de servicios en la detección del cumplimiento de objetivos de aprendizaje, o en el descubrimiento de conocimiento sobre el uso de los usuarios dentro del entorno de aprendizaje.

Palabras clave—Arquitectura software; *cloud*; indicadores de uso; análisis de uso; eLearning; Mundos Virtuales; *Second Life*

I. INTRODUCCIÓN

El análisis de la interacción de los usuarios con los entornos eLearning puede proporcionar conocimiento acerca de aspectos desconocidos de los usuarios. Entre este conocimiento, es posible destacar el descubrimiento de distintas habilidades, conocer la velocidad de adquisición de conocimientos, los deseos o gustos de los usuarios, sus hábitos de comportamiento, los patrones de uso que siguen, etc. [1-4]. Estas características pueden proporcionar al profesor un nuevo punto de vista acerca de cómo reaccionan, sienten y utilizan los alumnos ciertos ecosistemas de aprendizaje, de modo que pueda ser usado este conocimiento adquirido para mejorar los entornos de interacción, establecer perfiles entre los usuarios en función de

sus características (tanto desde un punto de vista personal, como en grupos), o simplemente para detectar aquellas áreas de los ecosistemas eLearning que deban ser mejoradas o adaptadas a los distintos tipos de usuarios que pueden utilizarlas. En el caso de este estudio, se plantea el uso de una arquitectura software basada en servicios y desplegada en *cloud* que sea capaz de medir y analizar datos extraídos de un entorno eLearning, proporcionando a los profesores responsables de dicho entorno eLearning la posibilidad de conocer las distintas características de sus alumnos respecto a una actividad de aprendizaje, pudiendo así llegar a evaluar si es posible establecer relaciones entre ciertos indicadores de uso y la nota final obtenida en la práctica o si es posible conocer los patrones de uso de los alumnos a través de una herramienta de este tipo. El entorno eLearning utilizado para este trabajo de investigación es un laboratorio virtual dentro del Mundo Virtual privativo *Second Life* [5]. Este laboratorio, conocido como *Usalpharma Lab*, simula las instalaciones, equipamiento, documentación, y demás elementos de un laboratorio que cumple con la normativa GLP (*Good Laboratory Practices*) de la industria farmacéutica [6]. La simulación virtual de un laboratorio de estas características se debe a la imposibilidad económica de construir y mantener una infraestructura de este tipo dentro de una Universidad, así como ofrecer la posibilidad de proporcionar una experiencia inmersiva y realista a alumnos de titulaciones relacionadas con el área de Farmacia en un contexto real de trabajo como el que se da en la industria. Este laboratorio, construido por personal relacionado con el Departamento de Farmacia y Tecnología Farmacéutica de la Universidad de Salamanca, lleva siendo usado desde 2011 para realizar actividades de formación con alumnos de posgrado en asignaturas relacionadas con la Calidad en la Industria Farmacéutica. La formación que reciben los alumnos en este escenario virtual se basa en el aprendizaje de metodologías,

normativas y procesos de auditorías en laboratorios de industria farmacéutica. Para ello, los alumnos actúan dentro de las instalaciones virtuales como auditores externos que se encargan de evaluar el cumplimiento del laboratorio de estas medidas GLP. En el caso que se presenta en este trabajo, se muestra cómo se ha incluido una arquitectura software que monitoriza la interacción de los usuarios dentro del laboratorio virtual [7], ofreciendo información a su vez sobre la consecución de los objetivos marcados en la práctica (número de errores detectados, tiempo usado, etc.) [8]. De este modo los alumnos pueden acceder al laboratorio virtual cuando ellos deseen (dentro de un plazo acordado con el profesor), realizando una inspección totalmente autónoma (sin presencia del docente), y sin limitación de número de accesos, hora o lugar de conexión. Una vez finalizado el plazo acordado para la inspección del laboratorio virtual, los alumnos entregan un informe que indica las deficiencias o puntos de la normativa que se incumplen, para que el profesor evalúe la práctica. En el caso descrito en este artículo, se trata del primer año de implantación de esta arquitectura software, por lo que las interacciones monitorizadas por el sistema no se han tenido en cuenta para evaluar la actividad, aunque puedan suponer una fuente de conocimiento acerca de la actividad de los usuarios dentro del laboratorio. Esta experiencia piloto fue llevada a cabo durante el mes de Marzo del 2014, y en ella participaron 9 alumnos del Máster Universitario en Evaluación y Desarrollo de Medicamentos de la Universidad de Salamanca, concretamente dentro de la asignatura de Garantía de Calidad en el Laboratorio de Análisis en la Industria Farmacéutica. Esta experiencia es la que se tomará como base para probar los conceptos presentados en este trabajo de investigación. Por ello, y como objetivo principal de este trabajo de investigación, se plantea utilizar una arquitectura software basada en servicios para medir la interacción de los usuarios dentro del laboratorio virtual que se encuentra dentro del Mundo Virtual Second Life, analizando distintas características de uso del mismo, observando finalmente si esta interacción medida tiene algún tipo de relación con la nota final obtenida en dicha práctica.

Este trabajo de investigación se divide en las siguientes secciones: esta primera (Introducción) trata de presentar al lector el contexto de la investigación, así como los rasgos básicos del caso de estudio y sus particularidades. Por su parte, la segunda sección (Materiales y Métodos) describe qué datos se miden dentro del laboratorio virtual y qué procesos se realizan para conseguir los objetivos marcados. A continuación, la tercera sección (Resultados y Discusión) presenta los resultados obtenidos en la etapa de análisis y explica cómo estos resultados pueden ayudar a determinar cierto conocimiento o son válidas para evaluar la actividad en un futuro. Por último, se incluye una sección de Conclusiones, donde se exponrán los conceptos y apartados más relevantes del trabajo.

II. MATERIALES Y MÉTODOS

A. Materiales

Para realizar este trabajo de investigación, se ha hecho uso de una arquitectura software basada en servicios implementada durante el año 2014 por varios miembros del equipo de este trabajo de investigación [7]. Esta arquitectura se basa en capas, de modo que desacopla sus servicios en niveles específicos de funcionalidad. Entre estas capas destacan la de recuperación de datos, la capa de persistencia, la capa de analítica, o la capa de provisión de datos a usuarios u otros servicios. Esta arquitectura, es capaz de recoger la interacción de los usuarios dentro de un entorno virtual, y componer una descripción *rica* (basada en una adaptación del protocolo RDF [9]) de la interacción ocurrida, para después poder almacenarla en la capa de persistencia (basada en bases de datos documentales NoSQL como MongoDB [10]), y finalmente proporcionar la posibilidad de establecer un análisis de datos de algún tipo acerca de los objetivos alcanzados por los alumnos y mostrarlos a través de vistas web a los profesores que desean conocer lo que ocurre dentro del ecosistema de aprendizaje usado (Second Life) [8]. Los factores relacionados con la interacción o el uso por parte de los usuarios sobre el laboratorio 3D que la plataforma es capaz de analizar o determinar, y que pueden hacer posible descubrimiento de la relación entre el uso y el rendimiento académico son los siguientes: tiempo total de uso del sistema, tiempo medio de sesión, número de sesiones, número de interacciones recogidas (número de clics).

- Tiempo total de uso: Suma de todos los tiempos de sesión que ha realizado el usuario dentro del sistema como conjunto. Este tiempo nos da una orientación real del nivel de *engagement*, dedicación, o compromiso que ha tenido el alumno respecto a la actividad realizada en las instalaciones 3D. En la subsección siguiente (Métodos) se explicará detalladamente cómo se ha llevado a cabo la estimación acerca del tiempo, ya que como Second Life se trata de un sistema cerrado, no es acceder a ningún tipo de archivo *log* del sistema, por lo cual es necesario estimarlo de otras formas.
- Tiempo medio de sesión: El tiempo medio de sesión puede ser un indicador fiable de la dedicación media del usuario por conexión, es decir, permite conocer si el usuario prefiere realizar visitas cortas (en cuanto a tiempo) al sistema pudiendo realizar acciones de auditoría pequeñas o sobre ciertos objetivos concretos, o si prefiere realizar acciones de una duración media mayor para cubrir más tiempo de análisis en cada una de ellas y dedicar más tiempo a la búsqueda de detalles en profundidad sobre las instalaciones.
- Número de sesiones: El número de sesiones puede ser un indicador sobre la fidelización (si un usuario se conecta

frecuentemente al escenario de aprendizaje, es posible que se sienta atraído de algún modo por él) y la evolución del alumno frente al entorno virtual. Este dato de número de sesiones, también puede indicar si el alumno necesita mayor o menor número de conexiones al Mundo Virtual para poder realizar sus actividades de auditoría de forma satisfactoria.

- Número de interacciones recogidas (número de clics): Esta medida puede ayudar a comprender la cantidad de opciones barajadas por el alumno para encontrar las deficiencias o no conformidades del laboratorio respecto de la normativa GLP (pudiendo encontrar casos de búsqueda indiscriminada de errores, u otros que denotan una búsqueda ordenada y razonada). También nos puede dar una medida de la actividad o inactividad del alumno dentro del laboratorio, si es capaz de manejarse con soltura o no dentro del mismo (por ejemplo estableciendo ratios de clics por segundo, o clics por sesión).

Estas medidas, junto con otras que puedan extraerse a partir de las mismas pueden permitir a los docentes conocer el uso del Mundo Virtual por parte de los alumnos y observar si este conocimiento adquirido tiene relación o no con la calificación final obtenida en la actividad de aprendizaje.

B. Métodos

A continuación se detalla cómo se ha llevado a cabo los distintos procesos que permiten a los profesores tener conocimiento acerca de lo que ocurre dentro del Mundo Virtual:

- Recogida de datos: Como se ha indicado previamente, conocer cualquier dato relacionado con la interacción dentro del Mundo Virtual Second Life es una tarea compleja. En el caso que nos ocupa, la interacción del usuario propiamente dicha, se ha realizado una interconexión (vía protocolo http) entre el entorno virtual y la arquitectura software, de modo que sea posible monitorizar a través de la capa de recogida de datos de la arquitectura. Para llevar a cabo este cometido, la recogida de datos del Mundo Virtual se organiza mediante una adaptación del protocolo RDF, de modo que todo dato recogido se expresa en el proceso de recolección en frases de tipo *sujeto (usuario) + verbo (acción que realiza) + predicado (objeto implicado, momento temporal, coordenadas virtuales, etc.)*. Estos datos enviados desde el Mundo Virtual quedan guardados en una capa de persistencia de datos dentro de la arquitectura (base de datos MongoDB), de modo que es posible acceder a estos datos analizarlos, pudiendo discriminar qué usuario ha realizado qué acciones, en qué momento temporal o sobre qué objeto concreto ha realizado alguna acción.
- Cálculo de factores de uso, sesiones y el tiempo de uso de los usuarios: Una vez almacenados los datos de

interacción, es posible detectar el número (técnicas de recuento, etc.) de acciones realizadas por cada usuario, el ratio de uso de cada objeto, el tipo de interacción del usuario y el entorno 3D (exploratoria, indiscriminada, ordenada, etc.). Además de conocer qué objetos se usan y por parte de qué usuarios, es posible conocer el tiempo de sesión y de conexión de los usuarios dentro del Mundo Virtual, ya que se ha implementado en la capa de servicios de analítica una técnica de estimación del tiempo. Concretamente, se sigue el mismo procedimiento que usa el servicio *Google Analytics* para estimar las sesiones y el tiempo de conexión de los usuarios en una página o portal web. Esta metodología se basa en el análisis de la consecución de eventos o acciones de un usuario dentro de un entorno digital, y determina, que el tiempo mínimo de conexión de un usuario dentro de un sistema, es aquel que transcurre entre la aparición del primer y el último evento de interacción. Para corregir errores como la posibilidad de utilizar el tiempo que un recurso digital se encuentre en reposo, es necesario contemplar el concepto de sesión de usuario. Una sesión de usuario se considera como aquel espacio de tiempo en el que un usuario está conectado de manera continuada a un sistema. Tomando el concepto de sesión, se puede calcular de una forma más fiable el tiempo de conexión al sistema, ya que solo se tiene en cuenta el tiempo de uso efectivo, no distanciado en demasía en el tiempo (diferencias inferiores a 30 minutos) [11, 12]. Por tanto, el tiempo de uso de los usuarios se considera finalmente como la suma del tiempo de las distintas sesiones, y el tiempo por sesión se calcula como el tiempo transcurrido entre el primer evento de interacción de la sesión y el último.

III. RESULTADOS Y DISCUSIÓN

Esta sección muestran como ejemplo del funcionamiento de la arquitectura los resultados obtenidos acerca de los indicadores de uso en 9 alumnos del Máster Universitario en Evaluación y Desarrollo de Medicamentos de la Universidad de Salamanca, comparándolos con las notas obtenidas. De igual forma, en esta sección se discute acerca de los resultados en términos de validez y utilidad, y del rendimiento de la arquitectura para realizar la recogida de datos y análisis de indicadores a través de ellos, y su posible para evaluar una actividad .

A. Resultados de la medición de los Indicadores de Uso

Una vez realizada la práctica y evaluada la actividad por parte de los profesores responsables de la asignatura, se ha realizado una medición a través de la arquitectura software de las distintas características de uso del entorno 3D, para poder conocer si la actividad monitorizada por los servicios de recogida de datos de la arquitectura se pueden relacionar con la

nota final de los alumnos, o por el contrario, para determinar si la interacción no tiene que ver tanto con la nota final (o por ejemplo, en este caso tiene mayor relevancia la teoría de la asignatura).

TABLA I. RESULTADOS DE LA MEDICIÓN DE LOS INDICADORES DE USO POR CADA ALUMNO Y NOTA OBTENIDA EN EL INFORME FINAL SOBRE LA PRÁCTICA Y LOS FALLOS DETECTADOS DENTRO DEL LABORATORIO VIRTUAL

Alumno	Número de interacciones recuperadas	Nº de sesiones	Tiempo medio por sesión (segundos)	Tiempo total de uso (segundos)	Nota del alumno (max 10 puntos)
1	1096	3	327	983	4,3
2	1533	9	190	1716	4,6
3	800	2	541	1082	4,8
4	1184	4	522	2090	7,5
5	831	5	390	1952	7,7
6	685	3	310	931	4,3
7	960	3	491	1475	5,5
8	1555	7	285	1995	7,3
9	1191	9	168	1517	5,5

En la Tabla 1 es posible observar los distintos datos medidos de cada alumno. Respecto a los resultados obtenidos de la medición de indicadores de uso, como se observa en la tabla, hay cierta disparidad en el tiempo de conexión de distintos alumnos, así como también se observa bastante diferencia entre unos alumnos y otros en el número de interacciones registradas. En general, es posible afirmar que hay un factor especialmente relacionado con la nota final, este factor es el tiempo dedicado a la realización de la práctica. En ese caso se observa que los alumnos (excepto uno, el alumno 2) que se han tomado más tiempo en la observación de los puntos de control que se deben revisar en el laboratorio son los que mejor nota han obtenido. En cuanto a la cantidad de interacciones recuperadas, no es posible determinar una relación directa con la nota, ya que se dan casos de alumnos que tienen registradas multitud de interacciones pero parece que no han interactuado con los objetos concretos que incluían algún fallo o deficiencia en cuanto a la normativa GPL. Del mismo modo, parece que el número de sesiones no tiene una relación directa con la calificación obtenida, ya que no se observa ninguna dependencia; por lo tanto el tiempo medio por sesión no tiene relación alguna tampoco.

B. Discusión acerca de los resultados

Tomando la comparación entre las mediciones realizadas y la nota obtenida por los alumnos, es posible afirmar que esta práctica por norma general, para realizarla de forma correcta

implica un tiempo prolongado de auditoría dentro del entorno virtual (posiblemente debido a la cantidad de detalles presentes en un laboratorio como el representado virtualmente). Esta afirmación se debe a que todos los alumnos, excepto uno, que habían aprobado la parte de redacción del informe final son los que más tiempo se habían tomado en realizar las actividades de inspección. A pesar de que haya tres de los indicadores que no tienen una relación directa con la calificación obtenida en el ejercicio práctico, hay que destacar que estos indicadores sirven al profesor para hacerse una idea del tipo de comportamiento que tienen los usuarios. No es igual que un usuario haya interactuado con muchos objetos (de manera acertada o no), que un alumno haya interactuado con muy pocos, ya que por ejemplo, la interacción con el medio puede considerarse como un síntoma de interés o curiosidad sobre el entorno virtual que le rodea, expresado en número de clics sobre objetos del entorno virtual. En cualquier caso, se puede considerar que la medición de los indicadores de uso mediante una arquitectura software, pueden indicar una relación entre los mismos y la nota final obtenida. Además pueden proporcionar una idea clara de cómo transcurre la actividad, permitiendo al profesor incluso controlar qué alumnos pueden estar arriesgándose a una evaluación final peor (alumnos con indicadores claramente inferiores al resto). En cualquier caso, se observa que la medición y análisis (aunque de un modo básico) de las variables que se pueden recoger de un entorno como el que se presenta en el caso de estudio pueden dar una información valiosa a los profesores o responsables de una actividad académica, y que la arquitectura software basada en capas de servicios que se ha presentado, tiene utilidad aún en su fase inicial, y proporciona valor a los docentes que la han usado.

IV. CONCLUSIONES

Este trabajo muestra cómo se puede aplicar una arquitectura software basada en servicios para el soporte y análisis de actividades de eLearning en un contexto como son los Mundos Virtuales. Este trabajo se ha centrado principalmente en cómo una arquitectura de este tipo puede incorporar funciones de medición y análisis de la información extraída de un contexto educativo digital para descubrir conocimiento acerca de cómo los usuarios realizan sus actividades de aprendizaje dentro de un ecosistema digital. Para ilustrarlo, se han medido tres indicadores de uso del caso de estudio presentado, a fin de observar si los docentes responsables de la actividad de aprendizaje pueden extraer conocimiento sobre sus alumnos. Así, este trabajo muestra cómo una arquitectura software puede medir y analizar diversos factores relacionados con el desempeño y realización de las actividades digitales, permitiendo al profesor conocer cómo los alumnos realizan actividades de aprendizaje, llegando a servir como apoyo a los docentes en cuanto a la detección de conductas o usos que

puedan llevar al alumno una mala realización una actividad de aprendizaje y poder aplicar medidas correctoras.

V. AGRADECIMIENTOS

Al Vicerrectorado de Política Académica de la Universidad de Salamanca por la financiación del proyecto de Innovación Docente 2013/201 bajo el que se ha podido desarrollar la arquitectura software usada en este trabajo.

VI. REFERENCIAS

- [1] C. Beer, K. Clark, and D. Jones, "Indicators of engagement," *Proceedings ascilite Sydney*, 2010.
- [2] B. Dalgarno and M. J. W. Lee, "What are the learning affordances of 3-D virtual environments?," *British Journal of Educational Technology*, vol. 41, pp. 10-32, 2010.
- [3] M. Fetscherin and C. Lattemann, "User acceptance of virtual worlds," *Journal of Electronic Commerce Research*, vol. 9, pp. 231-242, 2008.
- [4] K. Krause, "Understanding and promoting student engagement in university learning communities," *Paper presented as keynote address: Engaged, Inert or Otherwise Occupied*, 2005.
- [5] L. Lab. (2014, Feb 28, 2014). *Second Life*. Available: <http://secondlife.com>
- [6] C. Maderuelo, J. S. Pérez-Blanco, H. Zazo, J. M. Armenteros del Olmo, J. Cruz Benito, and A. Martín-Suárez, "Auditing Training Practice To Postgraduate Pharmacy Students In A Virtual World," presented at the 19th Annual Conference of the European Association of Faculties of Pharmacy (EAFP 2013), Ankara, Turkey, 2013.
- [7] F. J. García-Peñalvo, J. Cruz-Benito, C. Maderuelo, J. S. Pérez-Blanco, and A. Martín-Suárez, "Usalpharma: A Cloud-Based Architecture to Support Quality Assurance Training Processes in Health Area Using Virtual Worlds," *The Scientific World Journal*, vol. 2014, 2014.
- [8] J. Cruz-Benito, R. Therón, F. J. García Peñalvo, C. Maderuelo, J. S. Pérez-Blanco, H. Zazo, *et al.*, "Monitoring and feedback of Learning Processes in Virtual Worlds through analytics architectures: A real case," in *Sistemas y Tecnologías de Información. Actas de la 9ª Conferencia Ibérica de Sistemas y Tecnologías de Información*. vol. I Artículos, Á. Rocha, D. Fonseca, E. Redondo, L. P. Reis, and M. P. Cota, Eds., ed Barcelona, España, June, 18-21, 2014: AISTI (Asociación Ibérica de Sistemas y Tecnologías de Información), 2014, pp. 1126-1131.
- [9] G. Klyne, J. J. Carroll, and B. McBride, "Resource description framework (RDF): Concepts and abstract syntax," *W3C recommendation*, vol. 10, 2004.
- [10] MongoDB, "Top 5 Considerations When Evaluating NoSQL Databases," ed, 2013.
- [11] J. Cutroni. (2012, June, 20, 2014). *Understanding Google Analytics Time Calculations*. Available: <http://cutroni.com/blog/2012/02/29/understanding-google-analytics-time-calculations/>
- [12] Google. (2014, June, 20, 2014). *How Sessions are calculated in Analytics*. Available: <https://support.google.com/analytics/answer/2731565?hl=en>