



ESCUELA SUPERIOR DE INGENIERÍA

**Evaluación de competencias mediante la extracción de indicadores
procedentes de los registros de actividad de los entornos de aprendizaje
virtuales. Un mapeo sistemático de la literatura**

Trabajo de Investigación

Antonio Balderas Alberico

30 de noviembre de 2013



ESCUELA SUPERIOR DE INGENIERÍA

Itinerario Formativo de Doctorado 7009

Trabajo de Investigación

Evaluación de competencias mediante la extracción de indicadores procedentes de los registros de actividad de los entornos de aprendizaje virtuales. Un mapeo sistemático de la literatura

- Departamento: Ingeniería Informática
- Autor: Antonio Balderas Alberico
- Tutores: Manuel Palomo Duarte y Juan Manuel Dodero Beardo

Cádiz, 30 de noviembre de 2013

Índice general

1. Introducción	1
2. Metodología	3
2.1. Justificación	3
2.2. Preguntas de investigación	3
2.3. Protocolo de revisión	4
2.3.1. Motores de búsqueda	4
2.3.2. Términos de búsqueda	4
2.3.3. Criterios de selección	5
2.3.4. Esquema para la extracción de datos	6
3. Resultados	7
3.1. Selección de bibliografía	7
4. Conclusiones	11
Bibliografía y referencias	13

Indice de figuras

3.1. Distribución de publicaciones por años	9
3.2. Distribución de publicaciones por soporte	9

Indice de tablas

2.1. Resumen de búsqueda de bibliografía	5
3.1. Resumen de búsqueda de bibliografía	7
3.2. Resumen de selección de bibliografía	8
3.3. Resumen de selección de bibliografía	8
3.4. Resumen de selección de bibliografía	9

Capítulo 1

Introducción

El interés de la evaluación de competencias de aprendizaje ha cambiado del conocimiento a las competencias. Proyectos como el Tuning Educational Structures in Europe, apoyado por el Lifelong Learning Program de la Unión Europea, muestran la importancia de utilizar el concepto de competencia como base para los resultados de aprendizaje. Las competencias de aprendizaje son habilidades que un alumno ha de ser capaz de demostrar una vez que termina su formación. Estas competencias de aprendizaje se dividen en dos grupos: específicas y genéricas. Competencias específicas son aquellas relacionadas directamente con la utilización de conceptos, teorías o habilidades propias de un área en concreto, mientras que las competencias genéricas son habilidades, capacidades y conocimientos que cualquier estudiante debería desarrollar independientemente de su área de estudio [3]. Aunque obviamente sigue siendo muy importante el desarrollo del conocimiento específico de cada área de estudio, es un hecho que el tiempo y la atención también deben dedicarse al desarrollo de las competencias genéricas. También es importante reconocer la aplicación de dichas habilidades genéricas fuera del ámbito académico, ya que son cada vez más relevantes para la preparación de estudiantes para su futuro papel en la sociedad, en términos de empleabilidad y ciudadanía.

Desafortunadamente, evaluar ciertas competencias es a menudo una tarea subjetiva, tanto para el profesor como para los alumnos. A menos que una competencia genérica está directamente enlazada a una actividad específica, éstas son complejas de evaluar. Desarrollar un procedimiento detallado para la evaluación en el desempeño de los estudiantes en las competencias genéricas es una actividad compleja y que requiere mucho tiempo por los diferentes aspectos a tener en cuenta. Si los profesores a penas tienen tiempo suficiente durante el curso académico para cumplir su planificación y evaluar todas las tareas, exámenes o trabajos que los alumnos han tenido que realizar para demostrar la adquisición del conocimiento de la asignatura, más complicado será si además tienen que añadir tareas para evaluar las competencias genéricas. Por lo que, aunque un alumno haya superado una asignatura, no se puede garantizar que éste sea capaz de desempeñar las competencias genéricas que la profesión le va a demandar.

Esta actividad de evaluar se vuelve mucho menos escalable, siendo esta la justificación más importante de este trabajo, cuando los cursos están masificados. Un ejemplo de este tipo de cursos son los MOOC (Massive Open Online Course), cuya filosofía es la liberación del conocimiento, para que este llegue a un público más amplio, y para el que se suelen ofrecer plazas ilimitadas. Este tipo de curso, que también se caracteriza por ser de carácter abierto y gratuito, y con materiales accesibles de forma gratuita, presenta evidentes problemas de escalabilidad cuando la evaluación no está automatizada [5].

Como ya el lector habrá podido deducir, el ámbito de nuestro trabajo está relacionado con los entornos de aprendizaje virtual: LMS y VLE (LMS, Learning Management System y VLE, Virtual Learning Environment). Estos entornos están diseñados especialmente para tareas colaborativas ya que integran foros, wikis o forjas, y son ampliamente utilizados como un soporte para clases presenciales, haciendo más fácil la comunicación con los estudiantes y manteniendo las actividades y recursos para el tema [11, 7]. Además, son ampliamente utilizados hoy en día por los centros educativos a todos los niveles.

En los LMS cada archivo subido, cada acceso al sistema o cada comentario escrito en un foro por los estudiantes quedan registrados en el sistema. Algunos estudios muestran que se podrían analizar las situaciones que se dan en la aplicación de las tecnologías al aprendizaje (TEL, Technology Enhanced Learning) mediante la recopilación de los rastros de interacción producidos por estos entornos [1]. Sin embargo, la recopilación de la información almacenada acerca de las actividades del curso requiere filtrado antes de que pueda ser utilizado para el procesamiento de nivel superior [2]. Esta información recopilada podría ser utilizada como indicadores del trabajo de los estudiantes. Los indicadores son estados que determinan si se cumple un nivel para considerar la competencia como alcanzada. Se deben diseñar indicadores para que cada estudiante sea capaz de demostrar su competencia con independencia de su nota.

Este trabajo tiene como fin saber hasta qué punto la informática se ha ocupado de la extracción de este tipo de información. Qué procesos asistidos por ordenador y/o de manera automática se han utilizado para la evaluación de competencias. Mediante el estado del arte vamos a conocer hasta donde ha avanzado la investigación en este campo. Queremos saber qué se ha hecho y no, y cómo se ha hecho y dicho en torno a nuestro problema de investigación. Para ello vamos a realizar una revisión de la literatura, empezando por explicar la metodología en el siguiente capítulo, mostrando los resultados obtenidos en el tercero y las conclusiones en el último.

Capítulo 2

Metodología

Este estudio sistemático se basa en las directrices publicadas en la metodología propuesta por Kitchenham [6]. Esta metodología describe cómo se debe planificar, ejecutar y presentar los resultados de una revisión de la literatura en la ingeniería de software. Para este trabajo se ha utilizado la propuesta de Petersen [8].

2.1. Justificación

El motivo de este trabajo, tal como se describió en la introducción, es la localización de trabajos relacionados con la tecnología para la mejora del aprendizaje (TEL) cuyo fin sea la evaluación de competencias de manera automática. El contexto de esta investigación viene marcado por la problemática del “Open Data”. En la era de la Web 2.0, dispositivos como los gadgets de iGoogle se están convirtiendo cada vez más populares para la integración y personalización de contenidos Web de diferentes fuentes. Estas nuevas tecnologías de la Web se pueden aprovechar como portales de ciencia en términos de interfaz de usuario dinámica, que impulse la implementación de pasarelas con fines de educación avanzada, y así convirtiendo los datos en accesibles y reutilizables [10].

Los datos dentro de sistemas de tipo LMS o VLE, quedan registrados para el uso de estos sistemas, almacenando información de estudiantes, profesores, cursos, tareas, trabajos, etc., y relacionando y configurando todos estos elementos para ofrecer al usuario una experiencia de curso virtual. Haciendo todos estos datos almacenados accesibles, siguiendo la filosofía de “Open Data”, se podría sacar partido a los mismos con fines científicos. Y tal como se comentó en la introducción, para que esta evaluación de competencias sea escalable y más aún cuando el número de alumnos es importante, no queda otra alternativa que hacerlo de manera automatizada.

2.2. Preguntas de investigación

En este punto, hemos de plantearnos cuáles son las cuestiones que necesitamos resolver para llevar a cabo nuestra investigación. Necesitamos saber en qué punto se haya la comunidad científica en lo que se refiere al uso de los registros de los entornos de aprendizaje para evaluar las competencias genéricas. Para ello planteamos la primera pregunta: *¿Cuáles son las competencias que se han evaluado de forma automática o asistida por ordenador a partir del uso de los LMS?*

Además, necesitamos saber de qué forma lo hace, qué métodos y técnicas se utilizan para este fin. Para eso se plantean las preguntas segunda (*¿Qué métodos y técnicas se utilizan para evaluar competencias mediante el uso de un LMS?*) y tercera (*¿Con qué técnicas se pueden obtener evidencias (numéricas) objetivas del desarrollo de las siguientes competencias en un LMS?*).

Finalmente se necesita saber si los resultados que se obtengan de los registros de interacción de un sistema LMS realmente refleja el desempeño o no de los estudiantes en qué competencia, planteándose por tanto la cuarta pregunta: *¿Los resultados que se obtendrían de la evaluación de competencias genéricas de los alumnos mediante el uso de los registros de actividad de los LMS son indicadores fiables del desempeño de dichas competencias?*.

Preguntas de investigación:

1. ¿Cuáles son las competencias que se han evaluado de forma automática o asistida por ordenador a partir del uso de los LMS?
2. ¿Qué métodos y técnicas se utilizan para evaluar competencias mediante el uso de un LMS?
3. ¿Con qué técnicas se pueden obtener evidencias (numéricas) objetivas del desarrollo de las siguientes competencias en un LMS?
4. ¿Los resultados que se obtendrían de la evaluación de competencias genéricas de los alumnos mediante el uso de los registros de actividad de los LMS son indicadores fiables del desempeño de dichas competencias?

2.3. Protocolo de revisión

La definición del protocolo de revisión requiere la realización de una serie de acciones para obtener la bibliografía de nuestro mapeado. Comenzaremos indicando los motores de búsqueda que vamos a utilizar, qué términos de búsqueda utilizaremos en dichos motores y las herramientas de soporte a la revisión. Además se mostrarán qué criterios de inclusión de la bibliografía se siguen y el procedimiento de selección.

2.3.1. Motores de búsqueda

Una vez que las preguntas de la investigación se han establecido, hay que identificar con precisión la estrategia de búsqueda a seguir. En nuestro procedimiento, para encontrar la bibliografía, se realizarán consultas en las siguientes bibliotecas digitales:

- Wiley Online Library
- World Scientific Net
- Springer
- ACM Digital Library
- IEEE Digital Library (Xplore)
- Scopus

2.3.2. Términos de búsqueda

Existen muchos términos que pueden utilizarse para referirse a la evaluación de competencias genéricas de manera automatizada o asistida. Por la naturaleza de nuestro trabajo, debemos contemplar siempre en los términos de búsqueda los términos “assessment” y “generic skills” o “generic competences”. Además se añaden términos como “E-Learning”, “computer-assisted” o “mobile learning”. Sin embargo, incluir términos de este tipo reducen drásticamente el número de resultados obtenidos en la

búsqueda, no llegando a obtenerse bibliografía más significativa que si no se incluyen. Por tanto, a tenor de las pruebas se decide eliminar de la búsqueda ese tipo de términos. Por otro lado, sí se incluyen acrónimos de diferentes relacionados con las TEL, como son: “TEL“, “LMS“, “ICT“ (Information and communications technology) , “CBI“ (Content-Based Instruction). Y se descartan de la búsqueda términos como “ICE“ (Integrated Collaboration Environment) y “CSCL“ (Computer Supported Collaborative Learning). La combinación de los términos de búsqueda empleados en la investigación, así como a los motores de búsqueda que fueron aplicados cada una pueden comprobarse en la tabla 2.1.

SOURCE	SEARCH TERMS	SEARCH SCOPE
Wiley Online Library	assessment AND “generic competences“ OR “generic skills“ AND (TEL OR ICT OR CBI)	in All Fields
World Scientific Net	“generic competences“ OR “generic skills“ AND assessment	Anywhere in article
Springer	(“generic skills“ OR “generic competences“) AND students AND (TEL OR CBI OR ICT)	All fields (Including full text)
ACM Digital Library	(assessment and “generic skills“) and (TEL or LMS or ICT or CBI)	Any field (title, abstract, review)
ACM Digital Library	(assessment and “generic competences“) and (TEL or LMS or ICT or CBI)	Any field (title, abstract, review)
IEEE Digital Library (Xplore)	((TEL or LMS or ICT or CBI) AND (“generic skills“ OR “generic competences“)) AND assessment)	Full text and meta-data
Scopus	((TEL or LMS or ICT or CBI) AND (“generic skills“ OR “generic competences“)) AND assessment)	All fields (Including full text)

Tabla 2.1: Resumen de búsqueda de bibliografía

2.3.3. Criterios de selección

Para determinar si un trabajo debía formar parte de nuestra selección de estudios primarios se leyó tanto el título, como el resumen y las palabras clave. En ocasiones, esto no es suficiente, siendo necesario complementar la lectura anterior con una somera la lectura del artículo completo y más detallada de la introducción y las conclusiones. Nuestra búsqueda se centra en la localización de los trabajos en relación con los términos definidos en la sección 2.3.2. Por lo tanto, se realizó la proyección de los trabajos seleccionados utilizando los siguientes criterios de exclusión:

- Off Topic: trabajo no relacionado directamente con nuestra investigación. Son trabajos, que aún satisfaciendo los criterios de búsqueda porque de alguna forma se mencionan en el texto, después su contribución no está relacionada con la evaluación asistida de competencias genéricas.
- Unsupported Language: Trabajo escrito en un lenguaje diferente al inglés o español. La mayoría de los textos son en inglés, por lo que este criterio de descarte apenas es utilizado.
- Duplicated: Trabajos cuya contribución principal está recogida en otros trabajos ya incluidos.
- Unread: Trabajo que no ha podido ser leído. Son textos que no han sido leídos al no estar disponible para su lectura en las bibliotecas digitales a las que tenemos acceso desde la Universidad. Así mismo, aunque algún artículo ha podido ser encontrado gracias a Google Scholar o a la generosidad de sus autores, alguno ha quedado pendiente de lectura.

2.3.4. Esquema para la extracción de datos

Para la extracción de la información dividiremos los trabajos según el tipo de investigación que realizan.

Tipo de investigación

Según los autores existen diferentes enfoques para la investigación. Algunos de estos sistemas de clasificación son los propuestos por Wieringa [9] y Hevner [4]. Usamos el primero, ya que es el recomendado en el proceso de mapeo sistemático descrito por Petersen [8].

- Solución propuesta: se propone una solución para un problema; la solución puede ser innovadora o una extensión significativa de una técnica existente . Los posibles beneficios y la aplicabilidad de la solución se muestran por un pequeño ejemplo o una buena línea de argumentación.
- Investigación Validación: las técnicas investigados son nuevas y todavía no se han aplicado en la práctica. Estas técnicas pueden ser por ejemplo los experimentos, es decir, el trabajo realizado en el laboratorio.
- Evaluación de la Investigación: las técnicas se aplican en la práctica y se lleva a cabo una evaluación de la técnica. Se muestra cómo se implementa la técnica en la práctica (implementación de la solución) y cuáles son las consecuencias de la aplicación en términos de ventajas y desventajas (evaluación de implementación).
- Artículos de Experiencia: trabajos que explican qué y cómo algo se ha llevado a cabo en la práctica. Basado en la experiencia personal del autor.
- Artículos de opinión: estos trabajos expresan la opinión personal de alguien acerca de si una determinada técnica es buena o mala, o cómo se deben realizar las cosas. No se basan en metodologías de trabajo y de investigación relacionados.
- Trabajos filosóficos: estos trabajos esbozan una nueva forma de ver las cosas existentes, estructurando el campo en forma de una taxonomía o un marco conceptual.

Capítulo 3

Resultados

A continuación se muestran los resultados del estudio tal y como se describieron en la sección anterior.

3.1. Localización de la literatura

En la tabla 3.1 se muestran el número de publicaciones devueltas para cada búsqueda por las bibliotecas digitales utilizadas en este estudio. En total se recopilaron 467 trabajos, que fueron revisados para identificar si eran de utilidad para el estudio.

SOURCE	SEARCH TERMS	SEARCH SCOPE	RESULTS
Wiley Online Library	assessment AND “generic competences” OR “generic skills” AND (TEL OR ICT OR CBI)	in All Fields	140
World Scientific Net	“generic competences” OR “generic skills” AND assessment	Anywhere in article	20
Springer	(“generic skills” OR “generic competences”) AND students AND (TEL OR CBI OR ICT)	All fields (Including full text)	140
ACM Digital Library	(assessment and “generic skills”) and (TEL or LMS or ICT or CBI)	Any field (title, abstract, review)	57
ACM Digital Library	(assessment and “generic competences”) and (TEL or LMS or ICT or CBI)	Any field (title, abstract, review)	15
IEEE Digital Library (Xplore)	((TEL or LMS or ICT or CBI) AND (“generic skills” OR “generic competences”)) AND assessment)	Full text and metadata	48
Scopus	((TEL or LMS or ICT or CBI) AND (“generic skills” OR “generic competences”)) AND assessment)	All fields (Including full text)	47
TOTAL			467

Tabla 3.1: Resumen de búsqueda de bibliografía

3.2. Selección de trabajos

En la tabla 3.2 se muestran los resultados de la clasificación.

CRITERION	STUDIES	FRECUENCY
Included	64	13,70 %
Off Topic	374	80,09 %
Unsupported Language	1	0,21 %
Out Scope	2	0,43 %
Duplicated	20	4,28 %
Unread	6	1,28 %
Total	467	100 %

Tabla 3.2: Resumen de selección de bibliografía

3.3. Extracción de los datos

Se han seleccionado como estudio primario únicamente 64 artículos, un 13,7 % del total de los revisados. Casi la mayor parte de los seleccionados se pueden localizar en los últimos años, entre 2008 y 2013 (Tabla 3.3 y figura 3.1). Siendo las revistas el medio en el que más de este tipo de artículos se han publicado, tal y como se puede consultar en la tabla 3.4 y en la figura 3.2.

YEAR	RESULTS
1996	1
1997	1
1998	0
1999	1
2000	0
2001	0
2002	0
2003	1
2004	1
2005	2
2006	2
2007	4
2008	13
2009	4
2010	5
2011	6
2012	8
2013	14

Tabla 3.3: Resumen de selección de bibliografía



Figura 3.1: Distribución de publicaciones por años

PUBLICATION FORUM	RESULTS
Journal	36
Conference	15
Chapter	11

Tabla 3.4: Resumen de selección de bibliografía

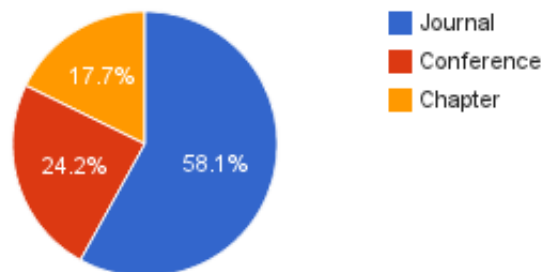


Figura 3.2: Distribución de publicaciones por soporte

3.4. Análisis y síntesis de los datos

....

Capítulo 4

Conclusiones

Conclusiones ...

Bibliografía

- [1] H. Chebil, J. Girardot, and C. Courtin. An ontology-based approach for sharing and analyzing learning trace corpora. In *Proceedings - IEEE 6th International Conference on Semantic Computing, ICSC 2012*, pages 101–108, 2012.
- [2] B. Florian, C. Glahn, H. Drachsler, M. Specht, and R. Fabregat Gesa. *Activity-based learner-models for learner monitoring and recommendations in moodle*, volume 6964 LNCS of *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer Berlin Heidelberg, 2011.
- [3] Julia Gonzalez and Wagenaar Robert. *Tuning Educational Structures in Europe - Final Report - Pilot Project - Phase 1*. University of Deusto, 2003.
- [4] Alan R. Hevner, Salvatore T. March, Jinsoo Park, and Sudha Ram. Design science in information systems research. *MIS Q.*, 28(1):75–105, March 2004.
- [5] D.H. Johnson. Teaching a “mooc:” experiences from the front line. In *Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE), 2013 IEEE*, pages 268–272, 2013.
- [6] Barbara Kitchenham, Rialette Pretorius, David Budgen, O. Pearl Brereton, Mark Turner, Mahmood Niazi, and Stephen Linkman. Systematic literature reviews in software engineering - a tertiary study. *Inf. Softw. Technol.*, 52(8):792–805, August 2010.
- [7] B. Munkhchimeg and S. Baigaltugs. Control possibility of students’ learning process through using learning management system. In *Strategic Technology (IFOST), 2013 8th International Forum on*, volume 2, pages 395–399, 2013.
- [8] Kai Petersen, Robert Feldt, Shahid Mujtaba, and Michael Mattsson. Systematic mapping studies in software engineering. In *Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering, EASE’08*, pages 68–77, 2008.
- [9] Roel Wieringa, Neil Maiden, Nancy Mead, and Colette Rolland. Requirements engineering paper classification and evaluation criteria: A proposal and a discussion. *Requir. Eng.*, 11(1):102–107, December 2005.
- [10] Wenjun Wu, M.E. Papka, and R. Stevens. Toward an opensocial life science gateway. In *Grid Computing Environments Workshop, 2008. GCE ’08*, pages 1–6, 2008.
- [11] A. Zafra, E. Gibaja, M. Luque, and S. Ventura. An evaluation of the effectiveness of e-learning system as support for traditional classes. In *Next Generation Web Services Practices (NWeSP), 2011 7th International Conference on*, pages 431–435, 2011.