



## **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 Estudio de Mapeo Sistemático**

Trabajo de Investigación

Antonio Balderas Alberico

12 de diciembre 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 Estudio de Mapeo Sistemático**

- Departamento: Ingeniería Informática
- Autor: Antonio Balderas Alberico
- Tutor: Manuel Palomo Duarte

Cádiz, 12 de diciembre de 2013



# Índice general

<b>1. Introducción</b>	<b>1</b>
<b>2. Metodología</b>	<b>3</b>
2.1. Justificación . . . . .	3
2.2. Preguntas de investigación . . . . .	4
2.3. Protocolo de revisión . . . . .	4
2.3.1. Motores de búsqueda . . . . .	5
2.3.2. Términos de búsqueda . . . . .	5
2.3.3. Criterios de selección . . . . .	6
2.3.4. Esquema para la extracción de datos . . . . .	6
2.3.5. Visualización y análisis de los datos . . . . .	8
<b>3. Resultados</b>	<b>9</b>
3.1. Localización de la literatura . . . . .	9
3.2. Extracción de los datos . . . . .	10
3.3. Categorización del estudio . . . . .	13
3.4. Esquema de clasificación . . . . .	16
3.4.1. Course's Learning Outcomes (CLO) and rubrics . . . . .	16
3.4.2. Peer and self eAssessment . . . . .	16
3.4.3. Game-Based Learning (GBL) . . . . .	17
3.4.4. eAssessment and reviews . . . . .	18
<b>4. Respuestas</b>	<b>23</b>
4.1. Respuestas a preguntas de investigación . . . . .	23
4.1.1. ¿Qué competencias se han evaluado de forma automática o asistida por ordenador a partir del uso de los entornos virtuales? . . . . .	23
4.1.2. ¿Qué herramientas o metodologías se utilizan para evaluar competencias mediante el uso de los entornos virtuales? . . . . .	24
4.1.3. ¿Con qué técnicas se pueden obtener evidencias (numéricas) objetivas del desarrollo de las siguientes competencias en un entorno virtual? . . . . .	24
4.1.4. ¿Los resultados que se obtendrían de la evaluación de competencias genéricas de los alumnos mediante el uso de los registros de actividad de los entornos virtuales son indicadores fiables del desempeño de dichas competencias? . . . . .	26
<b>5. Trabajo (presente y) futuro</b>	<b>27</b>
5.1. EvalCourse . . . . .	27
5.2. Proyectos en curso . . . . .	27
<b>6. Conclusiones</b>	<b>29</b>



# Indice de figuras

3.1. Distribución de las publicaciones por años . . . . .	11
3.2. Distribución de publicaciones según el medio en el que fueron publicados . . . . .	12
3.3. Distribución de publicaciones por tratamiento del problema . . . . .	13
3.4. Ámbito de trabajos distribuidos según tipo de investigación y según tipo de contribución.	15





# Indice de tablas

2.1. Resumen de búsqueda de bibliografía . . . . .	5
3.1. Bibliotecas digitales utilizadas, palabras de búsqueda utilizadas en cada uno y número de resultados obtenidos . . . . .	9
3.2. Clasificación de trabajos una vez aplicados los criterios de selección y exclusión . . . . .	10
3.3. Cantidad de trabajos publicados cada año . . . . .	10
3.4. Cantidad de trabajos según el medio en el que fueron publicados . . . . .	11
3.5. Distribución de las publicaciones . . . . .	12
3.6. Distribución de publicaciones por tratamiento del problema . . . . .	13
3.7. Distribución de publicaciones por tratamiento del problema . . . . .	21



# 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 [14]. 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 difíciles 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 apenas 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.

Dado que vivimos en un contexto donde las habilidades demandadas son cambiantes por la rápida transformación en una sociedad basada en el conocimiento, los sistemas de educación y formación en Europa, se es cada vez más conscientes de que los planes de estudio y con ellos las estrategias de evaluación deben centrarse en un enfoque más integral de las competencias genéricas y específicas. Las TIC ofrecen muchas oportunidades para el apoyo a los formatos de evaluación que pueden capturar habilidades complejas y competencias que son difíciles de evaluar [37]. Si los planes de estudio y los objetivos de aprendizaje han cambiado, también deberían hacerlo las prácticas de evaluación [35].

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 siempre disponibles las actividades y recursos para el tema [48, 29]. Además, son

ampliamente utilizados hoy en día por los centros educativos a todos los niveles.

Cuando el número de alumnos es elevado, se hace mucho menos escalable la evaluación para el profesor. 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 [17].

En los LMS, cada archivo subido, cada actividad realizada, cada acceso al sistema o cada comentario escrito en un foro por los estudiantes quedan registrados en el sistema. ¿Podríamos aprovechar la información almacenada sobre la interacción de los estudiantes con el sistema como indicadores del desempeño de ciertas competencias? Algunos estudios indican que así es, ya 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 [9]. Aunque se especifica que esta 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 [12].

El objetivo de este trabajo de investigación es realizar un análisis de la literatura para conocer hasta qué punto la ciencia se ha ocupado de la evaluación de competencias genéricas, prestando especial atención en aquellos trabajos que realicen dicha evaluación de manera automatizada aprovechando las prestaciones de las nuevas tecnologías. Nosotros pensamos que la información almacenada en la base de datos de un LMS puede ser aprovechada para extraer indicadores que nos den una medida del desempeño de los alumnos en ciertas competencias genéricas. ¿Hay algo de esto en la literatura? ¿Qué competencias pueden ser evaluadas de manera automática? Todas estas dudas, serán planteadas de manera formal en el siguiente capítulo, dónde se darán las indicaciones de la metodología seguida. A continuación se mostrarán los resultados, las respuestas a las preguntas, el trabajo futuro y por último las conclusiones y la bibliografía utilizada.

## Capítulo 2

# Metodología

Un Estudio de Mapeo Sistemático (SMS, Systematic Mapping Study), es una amplia revisión de los estudios primarios en un área específica cuyo objetivo es identificar alguna evidencia sobre el tema. Este estudio se basa en las directrices publicadas en la metodología propuesta por Kitchenham [19]. Esta metodología describe cómo se deben planificar, ejecutar y presentar los resultados de una revisión de la literatura en la ingeniería del software. Para este trabajo se ha utilizado la propuesta de Petersen [32].

Es muy importante seguir un procedimiento sistemático para llevar a cabo una revisión rigurosa de la literatura. En un primer momento, se trató de seguir la clasificación aportada por [38] para el ámbito de aplicación de la contribución. En él se plantea una división generacional:

- Primera generación: administración y calificación automatizada de pruebas convencionales
- Segunda generación: pruebas adaptativas
- Tercera generación: evaluación integrada continua
- Cuarta generación: retroalimentación tutorizada y personalizada

Sin embargo, la selección de la literatura realizada en este trabajo, no se adaptaba a esta categorización, por lo que finalmente los artículos fueron agrupados según el tipo de herramienta, modelo, paradigma o discusión que planteaban. La clasificación final fue:

- Course's Learning Outcomes (CLO) and rubrics: trabajos que evalúan las competencias genéricas de los alumnos a partir de rúbricas de los resultados del trabajo de los alumnos
- Peer and self eAssessment: trabajos que delegan la función de evaluar competencias genéricas en los compañeros o en la autoevaluación
- Game-Based Learning (GBL): trabajos que se basan en juegos para la evaluación de competencias genéricas
- eAssessment and reviews: resto de trabajos que agrupan propuestas para la evaluación de competencias genéricas mediante el uso de la tecnología

### 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 [47].

Los entornos de aprendizaje virtual (LMS ó VLE) almacenan información de estudiantes, profesores, cursos, tareas, trabajos, etc. Estos elementos se relacionan y configuran 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 entornos virtuales?*

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 herramientas o metodologías se utilizan para evaluar competencias mediante el uso de los entornos virtuales?*) y tercera (*¿Con qué técnicas se pueden obtener evidencias (numéricas) objetivas del desarrollo de las siguientes competencias en un entorno virtual?*).

Finalmente se necesita saber si los resultados que se obtengan de los registros de interacción de un sistema LMS realmente reflejan 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 entornos virtuales 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 entornos virtuales?*
2. *¿Qué herramientas o metodologías se utilizan para evaluar competencias mediante el uso de los entornos virtuales?*
3. *¿Con qué técnicas se pueden obtener evidencias (numéricas) objetivas del desarrollo de las siguientes competencias en un entorno virtual?*
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 entornos virtuales 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 estudio. 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) y 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 las palabras de búsqueda los términos “assessment” y “generic skills” o “generic competences”. Realizar la búsqueda por el término “Assessment of generic skills” o “assessing generic skills” nos planteaba la primera problemática, y es que el número de artículos devueltos era muy reducido. Sin embargo, debilitar la búsqueda con términos como “generic competences” o “generic skills” junto con la palabra “assessment” daba un número de resultados muy elevado. En primera instancia se probó añadiendo términos como “E-Learning”, “computer-assisted” o “mobile learning”. Sin embargo, incluir términos de este tipo reducían también 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 entornos virtuales relacionados con las TEL, como son: “TEL”, “LMS”, “ICT” (Information and communications technology), “CBI” (Content-Based Instruction). Y tras varias pruebas, se descartan también 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 en 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 que, habiendo sido obtenidos en el proceso de búsqueda anterior, vayan en línea con nuestro estudio y puedan ayudarnos a resolver las preguntas de investigación. Para ello, 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

Esta clasificación hace referencia al tipo de trabajo de investigación llevado a cabo por el/los investigador/es. Según los autores existen diferentes enfoques para la investigación. Algunos de estos sistemas de clasificación son los propuestos por Wieringa [46] y Hevner [16]. Usamos el primero, ya que es el recomendado en el proceso de mapeo sistemático descrito por Petersen [32].

- Solución propuesta (*proposal of solution*): 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 demuestran por un pequeño ejemplo o una buena línea de argumentación.
- Validación de investigación (*validation research*): las técnicas investigadas son nuevas y todavía no se han aplicado en la práctica. Estas técnicas podrían ser por ejemplo los experimentos, es decir, el trabajo realizado en un laboratorio.
- Evaluación de la Investigación (*evaluation research*): 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 (*experience papers*): 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 (*opinion papers*): 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 relacionadas.
- Trabajos filosóficos (*philosophical papers*): estos trabajos esbozan una nueva forma de ver las cosas existentes, estructurando el campo en forma de una taxonomía o un marco conceptual.

### Tipo de contribución

En este apartado se clasifican los trabajos según el tipo de contribución que realizan estos al ámbito en el que se desarrollan. Una vez realizado el estudio sistemático de la literatura y habiendo seleccionado los artículos, se realiza una clasificación en base a la aportación de éstos. El uso de algunos términos puede ser confuso, debido a la interpretación que hace el autor del mismo. Algunos de estos términos son framework, modelo, estrategia, proceso, procedimiento, método o metodología. Nuestra clasificación es la siguiente:

- Modelo (*model*): Es una representación de procesos, modelos o sistemas pertenecientes a un supra-sistema, cuyo fin es el análisis de interacción de ellos para mantener una relación flexible que les permita cumplir su función particular y cumplir la función de dicho supra-sistema.
- Proceso (*process*): contempla aquellos trabajos cuya contribución sea descrita por los autores como una secuencia de pasos.
- Herramienta (*tool*): se utiliza para los artículos que presentan un software independiente o una extensión de algún otro programa.
- Framework (*framework*): Aquí se consideran aquellos trabajos que contribuyen con una combinación de los elementos anteriores (es decir, con un modelo, un proceso y una herramienta).
- Técnica (*technique*): Un procedimiento utilizado para llevar a cabo una actividad o tarea específica. Podría venir acompañado de una herramienta de apoyo.

### Ámbito de aplicación de la investigación

Además de las clasificaciones anteriores, es necesario recoger más información acerca de los conceptos que representan la contribución de la investigación. Para ello se recoge información sobre el ámbito de la evaluación de competencias sobre el que se aplica cada contribución. Una vez recogida esta información, se agrupan según sus similitudes, quedando finalmente la siguiente clasificación:

- Resultados de aprendizaje del curso y rúbricas (*Course's Learning Outcomes (CLO) and rubrics*): Los resultados de aprendizaje del curso se evalúan mediante rúbricas o plantillas de evaluación en función del rendimiento de los alumnos. Esto proporciona al docente un indicador de sus logros de aprendizaje de cada alumno. Las rúbricas pueden estar o no en soporte informático, pero generalmente no aprovechan la tecnología para automatizar tareas.
- Autoevaluación y evaluación entre iguales (*peer and self eAssessment*): Uno de los problemas con los que se encuentran los profesores es la escalabilidad de la tarea de evaluación de competencias cuando el grupo de alumnos es grande. Hay un gran conjunto de trabajos, que aunque se apoyen en la tecnología para realizar alguna actividad, tienen el problema de que la evaluación ha de ser manual. En estos casos, mediante la autoevaluación o evaluación entre iguales los estudiantes se evalúan. De esta manera no sólo descargan de trabajo al profesor haciendo esta evaluación, sino que además se fomenta la capacidad crítica y de análisis del alumno.

- Aprendizaje basado en juegos (*Game-Based Learning (GBL)*): El aprendizaje basado en juegos se sirve de juegos que están diseñados expresamente para enseñar al usuario acerca de ciertos temas, ampliar conceptos o reforzar el desarrollo o aprendizaje de una habilidad mientras juegan.
- E-Evaluación y revisiones (*eAssessment and reviews*): Trabajos en los que se obtienen indicadores del desempeño de estudiantes en una o varias competencias de manera automática mediante el uso de algún software. Además se muestran otros trabajos sobre la situación actual en la evaluación de competencias genéricas, su importancia actual y sobre un conjunto de técnicas, metodologías o herramientas que se han desarrollado y utilizado.

### **2.3.5. Visualización y análisis de los datos**

Tras obtener los estudios primarios, hay una etapa de análisis, donde se resumen los datos extraídos para así responder a las preguntas de investigación planteadas. El análisis de los resultados se centra en el estudio de las publicaciones para cada categoría y por lo tanto, la determinación de lo bien que está cubierta cada categoría. Esta información generalmente se resume en los cuadros o mediante gráficos de barras. Otro método utilizado en nuestro estudio es la combinación de diferentes categorías (por ejemplo, el ámbito de investigación contra el tipo contribución) y mostrarlos en un mapa sistemático en la forma de un gráfico de burbujas. En el siguiente capítulo se mostrarán los resultados obtenidos.

## Capítulo 3

# Resultados

A continuación se muestran los resultados del estudio tal y como se describieron en la sección anterior. Comienza el capítulo con la localización de los estudios primarios, para continuar con la extracción de los datos de estudio, mostrándose varios gráficos o tablas que justifican la información mostrada. Finalmente se categorizan los estudios y se muestra el esquema de clasificación resultante.

### 3.1. Localización de la literatura

En la tabla 3.1 se muestran las búsquedas realizadas en las bibliotecas digitales más importantes en ciencias de la computación, los términos de búsqueda utilizados y el número de documentos obtenidos. En cada biblioteca, se utilizaron los formularios de búsqueda avanzada.

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)	142
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			468

Tabla 3.1: Bibliotecas digitales utilizadas, palabras de búsqueda utilizadas en cada uno y número de resultados obtenidos

En total se recopilaron 468 trabajos para ser revisados. El número de estudios primarios resultante (después de aplicar criterios de selección y exclusión) fue de sólo 32 trabajos (casi un 7 % del total de trabajos recopilados). Aunque hay muchos trabajos que tratan las competencias genéricas desde diferentes perspectivas, son muy pocos los que abordan su evaluación desde el punto de vista de la tecnología. De ahí estos resultados, cuya primera y optimista interpretación es que pudiera haber un nicho de investigación. Los resultados de esta clasificación pueden verse en la tabla 3.2.

CRITERION	STUDIES	FREQUENCY
Included	32	6,84 %
Off Topic	407	86,97 %
Unsupported Language	1	0,21 %
Duplicated	20	4,27 %
Unread	8	1,71 %
Total	468	100 %

Tabla 3.2: Clasificación de trabajos una vez aplicados los criterios de selección y exclusión

## 3.2. Extracción de los datos

Aunque hace varios años desde que las tecnologías entraron a formar parte de la vida académica, no es hasta 2010, con lo que la Comisión Europea llama la tercera generación de herramientas, cuando se comienzan a integrar la evaluación en las herramientas de aprendizaje, y conceptos como “Data Mining and analysis”, “Behavioural tracking” and “Learning analytics” comienzan a asomar [37]. Con esta introducción, damos sentido a la distribución de la producción de la selección primaria a lo largo de los años, que puede verse tanto en la tabla 3.3 como en la figura 3.1. Casi la mayor parte de los seleccionados se pueden localizar en los últimos años, entre 2011 y 2013.

YEAR	RESULTS
2003	0
2004	0
2005	0
2006	2
2007	1
2008	6
2009	1
2010	2
2011	5
2012	5
2013	10

Tabla 3.3: Cantidad de trabajos publicados cada año



Figura 3.1: Distribución de las publicaciones por años

La revista es 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 con un 53,1 % del total. Esta información se complementa con la distribución de las publicaciones según el foro en el que han sido publicados y que se muestra en la tabla 3.5. En este se puede comprobar como la mayor parte de las publicaciones están relacionadas con la ingeniería y la educación: World Scientific and Engineering Academy and Society Conferences (WSEAS), IEEE Global Engineering Education Conference (EDUCON), European Journal of Education o Revista Iberoamericana de Tecnologías del Aprendizaje (RITA), entre otras, son las publicaciones que más trabajos han aportado a nuestro estudio.

PUBLICATION TYPE	RESULTS
Journal	18
Conference	10
Chapter	5

Tabla 3.4: Cantidad de trabajos según el medio en el que fueron publicados

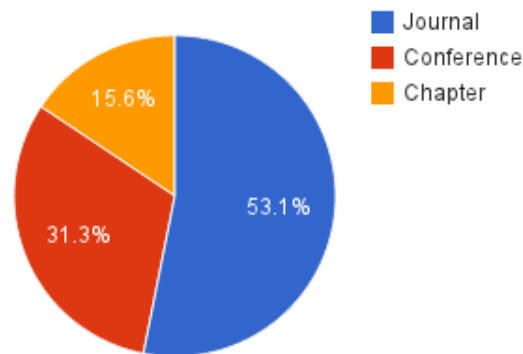


Figura 3.2: Distribución de publicaciones según el medio en el que fueron publicados

PUBLICATION FORUM	PAPERS
World Scientific and Engineering Academy and Society Conferences	4
IEEE Global Engineering Education Conference	2
Australian Society for Computers in Learning in Tertiary Education	2
European Journal of Education	2
Revista Iberoamericana de Tecnologías del Aprendizaje	2
Assessment & Evaluation in Higher Education	1
Australasian Journal of Educational Technology	1
Conference on Software Engineering Education and Training	1
Competency-based Language Teaching in Higher Education	1
Computers in Human Behavior	1
Computing Colombian Conference	1
Decision Support Systems	1
Game-based learning in higher education and lifelong learning: bridging the gap between theory and practice	1
Human Factors and Ergonomics in Manufacturing & Service Industries	1
International Conference on Advanced Learning Technologies	1
International Journal of Learning Technology	1
Journal of Computer Assisted Learning	1
Medical Education	1
Revista de Educación	1
The Internet and Higher Education	1
Ubiquitous and Mobile Learning in the Digital Age	1

Tabla 3.5: Distribución de las publicaciones

### 3.3. Categorización del estudio

Una vez revisados todos los artículos, se han extraído unas características o categorías comunes a la tipología de los trabajos. Todos los trabajos seleccionados hacen uso de algún tipo de software o metodología para evaluar algún tipo de competencia genérica. Pero ningún trabajo utiliza un enfoque como el que se propone en la introducción de este capítulo, es decir, aprovechando los registros de interacción de los estudiantes con el LMS como indicadores del desempeño de las competencias genéricas. Encontramos trabajos que se apoyan en la tecnología para el tratamiento o evaluación de las competencias, pero que terminan delegando parte de esta evaluación en el alumnado, ya sea mediante autoevaluación o evaluación entre iguales. Otros trabajos se basan en videojuegos o en las redes sociales para evaluar alguna competencia, mientras que otros desarrollan algún tipo de software o técnica. Finalmente hay algunos trabajos que simplemente detectan en su entorno la necesidad de la evaluación de las competencias de manera automática porque su forma de hacerlo les ocasiona una serie de problemas o desventajas con respecto a otro método que proponen o demandan. Además se han encontrado algunas revisiones sobre la literatura relacionadas que también serán tratadas a parte. En la tabla 3.6 se puede ver la distribución de las publicaciones, apoyadas gráficamente en la figura 3.3.

PUBLICATION ISSUE	PAPERS
Course's Learning Outcomes (CLO) and rubrics	7
Peer and self eAssessment	14
Game-Based Learning (GBL)	6
eAssessment and reviews	5

Tabla 3.6: Distribución de publicaciones por tratamiento del problema

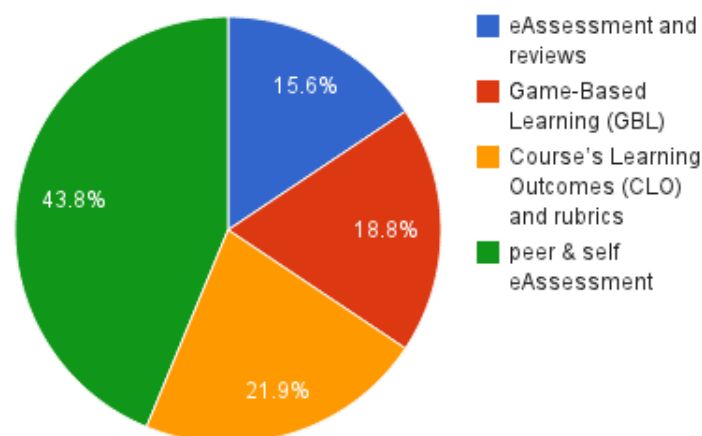


Figura 3.3: Distribución de publicaciones por tratamiento del problema

La distribución de los trabajos clasificados según su ámbito y su tipo por un lado, y según su ámbito y su contribución por otro, se puede ver en la figura 3.4. En esta figura se puede ver el grado de inmadurez de esta disciplina. El primer indicador para corroborar esta interpretación se puede encontrar en el gráfico de la derecha. En este se muestran la relación de trabajos a partir de su ámbito y de su tipo de investigación, como no se han hallado trabajos que nos narren la experiencia real (*Experience papers*) de un trabajo o trabajos en los que el autor reflexione sobre un campo concreto (*Philosophical papers*); y muy pocos trabajos de opinión (*Opinion papers*). También en el gráfico de la derecha vemos que el número de trabajos en cada campo está muy distribuido, y que la única propuesta que levemente destaca es la propuesta (*Proposal of solution*) de trabajos de autoevaluación o evaluación entre iguales (*Peer & self assessment*). Esta distribución de trabajos es otra evidencia del precoz grado de desarrollo de esta disciplina.

Si nos centramos en el lado izquierdo de la figura, dónde se relaciona el ámbito del trabajo con el tipo de contribución, vemos que los trabajos están también sumamente repartidos, aunque esta vez con participación de todos los elementos que participan en la matriz. El hecho de que participen todas las intersecciones y de que no haya ninguna tendencia que destaque es un indicador de que la investigación en este área aún no ha dado con un planteamiento en el que asentarse.



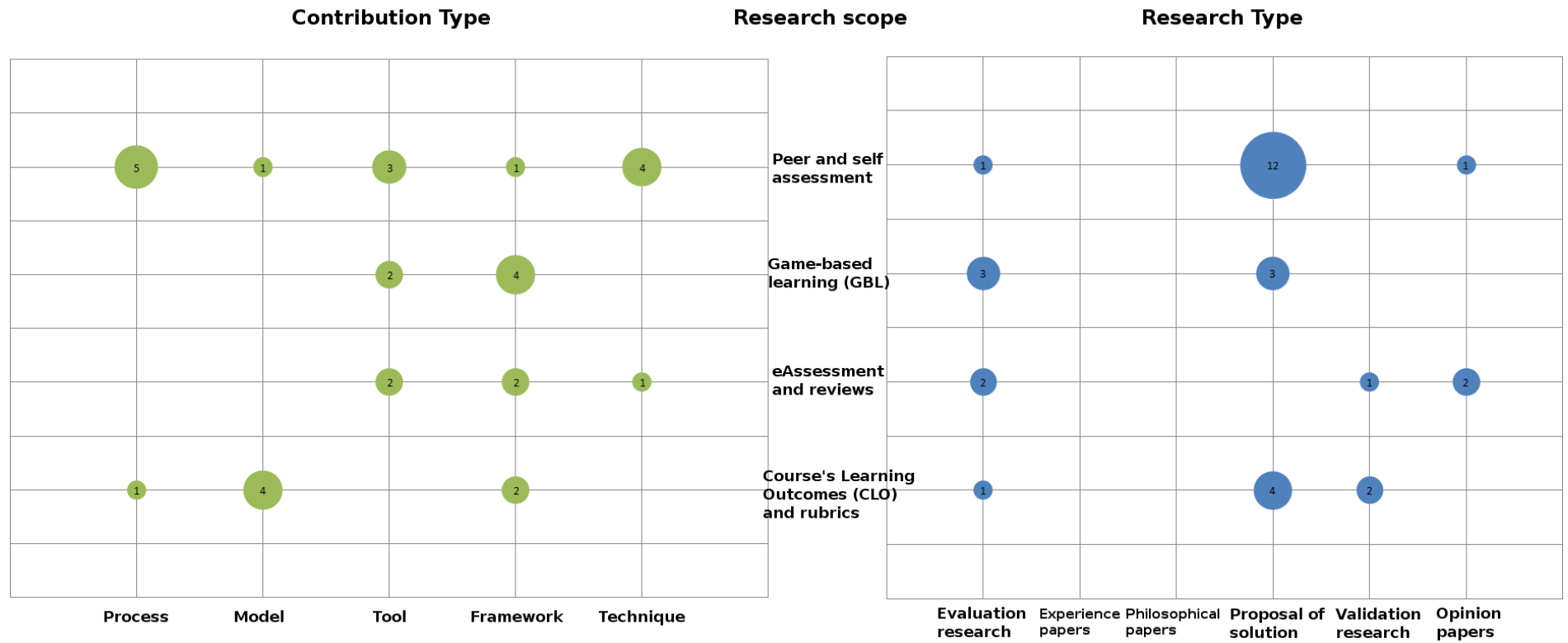


Figura 3.4: Ámbito de trabajos distribuidos según tipo de investigación y según tipo de contribución.

### 3.4. Esquema de clasificación

En los apartados siguientes, se presentan los resultados del estudio para cada área de investigación. El listado de trabajos se muestra en la tabla 3.7.

#### 3.4.1. Course's Learning Outcomes (CLO) and rubrics

Hoy en día, las universidades y las empresas van cada vez más de la mano en cuanto hablamos de competencias. Y así ocurre en numerosos artículos de los encontrados en este mapeado cuando se trata de buscar el desarrollo y evaluación de competencias. Y esto es así porque los empresarios buscan profesionales capaces de renovar constantemente sus conocimientos y competencias. Ser capaces de desarrollar y evaluar competencias se convierte, por tanto, en fundamental dentro del nuevo paradigma educativo que emerge del proceso de Bolonia. En este apartado se presentan trabajos que utilizan plantillas generales para las competencias genéricas como una herramienta útil que proporciona información sobre el desarrollo y la evaluación de cada habilidad, con diferentes tipos de matrices de valoración y las plantillas de evaluación. En [44] se introducen habilidades de empleo genéricos en los programas de los estudiantes de educación superior y son evaluadas mediante plantillas de evaluación, mientras que en [11] se echa en falta una evaluación de competencias genéricas y se proponen este tipo rúbricas para ello.

Siguiendo con el ámbito empresarial, nos topamos con el Consorcio RH-XML, una organización independiente y sin ánimo de lucro, dirigido por voluntarios dedicada al desarrollo y promoción de un conjunto estándar de especificaciones para que el comercio electrónico y la automatización de los intercambios de datos relacionadas con recursos humanos. En [39] proponen *HR-XML Competencies* con el fin de medir características que sirven como indicadores de competencias. Se pretende establecer unos valores estándares para integrarlos en los recursos del consorcio.

Por otro lado, cuando hablamos de *Course's Learning Outcomes (CLO)* nos referimos a objetivos de aprendizaje que describen con claridad las competencias que un estudiante debe poseer al completar el curso [41, 2, 18]. En [28, 27, 36] se implementan diversos modelos para mediante diferentes indicadores valorar el desempeño de los estudiantes en las competencias. En [26] partiendo de metodologías de aprendizaje social y colaborativo, utilizando las plataformas de gestión como Moodle, las redes sociales y Second Life, se presenta un trabajo que tiene como objetivo presentar un conjunto de estrategias de enseñanza, tareas y prácticas para desarrollar las competencias transversales necesarias para el estudio tanto éxito y futura integración profesional. Se evalúan las competencias de liderazgo, trabajo en equipo y razonamiento crítico mediante el uso de rúbricas y la corrección de varios profesores.

En resumen, se puede decir que estos trabajos tratan la evaluación de competencias mediante el uso de rúbricas, estableciéndose unos indicadores que marcarán en éstos si los alumnos o el personal a evaluar ha alcanzado el desempeño en éstas. Aunque pueden servir como referencia a la hora de evaluar competencias genéricas, estos sistemas no están integrados en las herramientas de evaluación, y aún estándolos, no liberan al docente de la carga que supone la corrección o revisión de dichas rúbricas.

#### 3.4.2. Peer and self eAssessment

En muchos trabajos se fomenta el uso de alguna herramienta de tipo *portfolio* o diario de aprendizaje, incluídas en los LMS actuales. Estos se han convertido en una ventana que proporciona información del trabajo del estudiante y permite conocer su forma de pensar [31, 13]. A posteriori son los compañeros o los propios estudiantes los que deben evaluar su evolución y analizar si han conseguido o mejorado ciertas competencias genéricas. Igualmente ocurre en [21], donde los estudiantes trabajan en grupo en un wiki y posteriormente evalúan las competencias de sus compañeros mediante el uso de una rúbrica. Otro tipo de estrategia con una idea parecida se utilizan en la evaluación entre iguales es la integración de los entornos de aprendizaje con las redes sociales [34, 24, 40]. Los profesores proporcionan rúbricas que

permiten a los estudiantes entender las expectativas de sus instructores. Los alumnos proporcionan información directa a sus compañeros acerca de lo que han aprendido y lo que todavía tienen que aprender. En segundo lugar, los compañeros o ellos mismos pueden usar las rúbricas para la auto-evaluación.

Una experiencia muy similar pero utilizando herramientas multimedia es la que se lleva a cabo en otro grupo de trabajos. Uno de ellos es el proyecto mostrado en [22], donde se realizan videoconferencias entre estudiantes de una universidad española y otra estadounidense. Así se pretende promover un aprendizaje autónomo y la mejora de la competencia de comunicación en un segundo idioma. Otro trabajo es el realizado en [23], donde mediante autograbación se evalúan ciertas competencias genéricas. Los alumnos son divididos en grupos por el profesor, se graban en vídeo, se cuelgan en el campus y se auto-corrigiendo viendo sus fallos y mejorándolos. Son dos experiencias en las que los alumnos corrigen sus propios trabajos y son capaces de valorar si han alcanzado o no ciertas competencias a partir de su evolución.

Otros artículos en los que se trabaja también con la autoevaluación son [10, 20, 25, 30, 9, 8]. Con este grupo de trabajos se soluciona el problema de escalabilidad con la autoevaluación o evaluación entre iguales. Sin embargo, la calificación depende en gran medida de la subjetividad del alumno y una revisión minuciosa del profesor seguiría siendo un trabajo poco escalable.

### 3.4.3. Game-Based Learning (GBL)

La aplicación de los conocimientos teóricos a la práctica es uno de los objetivos más importante y hacia esta dirección apuntan diversos métodos y enfoques que se han desarrollado. Un aprendizaje significativo incrustado en la experiencia es difícil de proporcionar dentro de los planes de estudios regulares. Sería ideal poder trabajar sobre situaciones de la vida real, y poder integrar este tipo de experiencias en los planes de estudios universitarios sin riesgos. En varios trabajos se muestra la integración de los entornos de trabajo auténticos con el fin de apoyar la aplicación práctica de los conocimientos teóricos dentro de los enfoques de enseñanza y aprendizaje en los cursos de postgrado y formación permanente. Estas experiencias se llevan a cabo con simulaciones y juegos [33, 7].

Los *juegos serios* ofrecen nuevos desafíos y oportunidades para el desarrollo de competencias. Este potencial implica cuestiones de investigación complejas, tales como la forma de evaluar las competencias sin perturbar el juego en sí al que se enfrenta el estudiante, así como la forma de diseñar el juego de modo que las competencias pueden ser desarrolladas por los alumnos. A fin de evaluar las competencias, es útil describirlas en términos de la conducta que se observa e interpreta mientras que un estudiante está involucrado en el juego. Sin embargo, la interpretación de estos indicadores de comportamiento en términos de competencias, y en particular, de las habilidades sociales, está mediada por factores contextuales [5]. En otro trabajo basado en un juego serio [15], se utiliza una aventura gráfica mediante la que los alumnos desarrollan y evalúan sus propias competencias de espíritu empresarial y resolución de problemas.

Otra herramienta de simulación es VIRBUS [42, 43]. VIRBUS un juego que simula un problema de la vida real y ayuda evaluar la competencia genérica de *resolución de problemas* mediante la salida esperada al problema y la encontrada por el alumno (o grupo de alumnos). El objetivo principal de la aplicación de esta herramienta es facilitar el aprendizaje autodirigido del estudiante. Además, los estudiantes obtienen retroalimentación, no sólo una evaluación de sus resultados de aprendizaje.

Los juegos ofrecen mecanismos para desarrollar y evaluar las competencias de sus estudiantes, y algo muy importante, es que los alumnos aprenden jugando. Aunque nos enfrentamos a varios problemas con respecto a nuestra propuesta. Por un lado, son herramientas que no están integradas en el sistema, por lo que conllevan consigo algún mecanismo manual de traslado de calificaciones. Además, los juegos suelen ir orientados a un reducido grupo de competencias, readaptar el juego para evaluar otras competencias sería algo muy costoso. Y tener varios juegos puede encaminarnos a un nuevo panorama de trabajo poco escalable para el profesor.

#### 3.4.4. eAssessment and reviews

Este grupo de artículos están más orientados a la tercera generación en las estrategias de evaluación (*eAssessment*), enfocados en minería de datos, seguimiento del comportamiento y análisis integrado de aprendizaje de los estudiantes en lo referente a su interacción con el entorno virtual [38, 37]. En los artículos encontrados se realiza una revisión del estado del arte actual y se plantean las bases para nuevas herramientas.

Por último quedan una serie de trabajos que evalúan competencias, y aunque su faceta tecnológica no ha podido ser desgranado con precisión, dada la ausencia de información en el trabajo, sí se puede vislumbrar que se ha apoyado en algún u otro aspecto de la tecnología para su desempeño. En [45] se combinan instrumentos para evaluar a los estudiantes como exámenes tradicionales, sesiones prácticas con ordenador y trabajos en grupos para evaluar la adquisición de competencias. En [1] se muestra Tripuscoid, una herramienta que, en primer lugar, permite a los estudiantes a identificar sus puntos fuertes y débiles y desarrollar estrategias personales para la mejora, en segundo lugar, proporciona a los profesores información adicional sobre los efectos de su entrada en las competencias de los estudiantes, y por último, proporciona información útil para gestión de la calidad de los programas de enseñanza, ya que puede detectar necesidades de formación de los nuevos estudiantes y ayudar a mejorar el contenido y diseño de los futuros programas académicos. En [6] se examinan algunos casos especiales para los que las nuevas tecnologías han permitido avances significativos (por ejemplo, las evaluaciones de los alumnos con necesidades educativas especiales, la evaluación de habilidades de colaboración y la consecución de grupo).

Aunque su aportación es interesante para este trabajo, los métodos presentados distan de las soluciones que necesitamos para tratar los problemas de escalabilidad y subjetividad.

REF	TITLE	RESEARCH SCOPE	RESERACH TYPE	CONTRIBUTION TYPE
[9]	An Ontology-Based Approach for Sharing and Analyzing Learning Trace Corpora	Peer & self eAssessment	Solution proposal	Tool
[27]	Appraisal of course learning outcomes using rasch measurement: a case study in information technology education	Course's Learning Outcomes (CLO) & rubrics	Evaluation research	Model
[44]	Assessing Transferable Generic Skills in Language Degrees	Course's Learning Outcomes (CLO) & rubrics	Solution proposal	Framework
[24]	Beyond marks and measurement: Developing dynamic and authentic forms of e-assessment	Peer & self eAssessment	Solution proposal	Technique
[40]	Can Web 2.0 technology assist college students in learning English writing? Integrating Facebook and peer assessment with blended learning	Peer & self eAssessment	Solution proposal	Technique
[33]	Challenges and opportunities in evaluating learning in serious games: a look at behavioural aspects	Game-Based Learning (GBL)	Solution proposal	Framework
[37]	Changing Assessment — Towards a New Assessment Paradigm Using ICT	eAssessment and reviews	Evaluation research	Framework
[1]	Competence Assessment in Higher Education: A Dynamic Approach	eAssessment and reviews	Validation research	Tool
[10]	Competence gaps in software personnel: A multi-organizational study	Peer & self eAssessment	Solution proposal	Process
[39]	Competence Models in Technology-Enhanced Competence-Based Learning	Course's Learning Outcomes (CLO) & rubrics	Solution proposal	Model
[13]	Cooperative learning and electronic group portfolio: tutoring tools, development of competences and assessment	Peer & self eAssessment	Solution proposal	Technique
[20]	Developing a diagnosis system of work-related capabilities for students: A computer-assisted assessment	Peer & self eAssessment	Solution proposal	Process
[45]	Developing Generic Competences in the European Higher Education Area: a proposal for teaching the principles of economics	eAssessment and reviews	Evaluation research	Framework
[42]	Developing virtual simulation game for authentic learning: realizing partnership between university and industry	Game-Based Learning (GBL)	Validation research	Framework

[38]	eAssessment for 21st century learning and skills	eAssessment and reviews	Opinion papers	Technique
[36]	Engineering students performance evaluation of generic skills measurement: ESPEGS model	Course's Learning Outcomes (CLO) & rubrics	Solution proposal	Model
[26]	Everything Matters: Development of Cross-Curricular Competences in Engineering Through Web 2.0 Social Objects	Course's Learning Outcomes (CLO) & rubrics	Solution proposal	Process
[25]	Explorations in metacognition: The design, development, and implementation of an online teamwork tracking environment	Peer & self eAssessment	Solution proposal	Tool
[5]	From Behavioral Indicators to Contextualized Competence Assessment	Game-Based Learning (GBL)	Validation research	Framework
[43]	Game-based learning in higher education and lifelong learning: bridging the gap between theory and practice	Game-Based Learning (GBL)	Solution proposal	Framework
[30]	Generic skills in medical education: developing the tools for successful lifelong learning	Peer & self eAssessment	Solution proposal	Tool
[11]	Generic Skills in Software Engineering Master Thesis Projects: Towards Rubric-Based Evaluation	Course's Learning Outcomes (CLO) & rubrics	Validation research	Framework
[22]	Innovation Network: Videoconferencing as a Resource in Teaching Support and Autonomous Learning	Peer & self eAssessment	Evaluation research	Framework
[34]	Measuring collaboration and creativity skills through rubrics: Experience from UTPL collaborative social networks course	Peer & self eAssessment	Solution proposal	Process
[28]	Outcome based education performance measurement: a Rasch-based longitudinal assessment model to measure information management courses LO's	Course's Learning Outcomes (CLO) & rubrics	Validation research	Model
[23]	Self-video recording for the integration and assessment of generic competencies	Peer & self eAssessment	Solution proposal	Process
[15]	Serious Games for the Development of Employment Oriented Competences	Game-Based Learning (GBL)	Validation research	Tool
[7]	SIMBA: A simulator for business education and research	Game-Based Learning (GBL)	Solution proposal	Tool
[6]	Technological Issues for Computer-Based Assessment	eAssessment and reviews	Opinion papers	Tool
[31]	The educational model at university and the use of new methodologies for teaching, learning and assessment	Peer & self eAssessment	Solution proposal	Technique

[8]	Towards a model for assessing competencies	Peer & self eAssessment	Opinion papers	Model
[21]	Using wikis to develop student teachers' learning, teaching, and assessment capabilities	Peer & self eAssessment	Solution proposal	Process

Tabla 3.7: Distribución de publicaciones por tratamiento del problema





## Capítulo 4

# Respuestas

En esta sección, se trata de dar respuesta a las preguntas de investigación.

### 4.1. Respuestas a preguntas de investigación

En los artículos de este trabajo se tratan diversas competencias. Unas se trabajan y evalúan, otras sólo se mencionan y otros evalúan competencias no reconocidas como tal. En los siguientes apartados, correspondientes a cada una de las preguntas de investigación, se han separado las respuestas por competencias, indicando qué artículos pueden servir como ejemplo para cada caso.

#### 4.1.1. ¿Qué competencias se han evaluado de forma automática o asistida por ordenador a partir del uso de los entornos virtuales?

Las competencias que se han evaluado en los artículos de este trabajo son las siguientes:

- Capacidad para comunicarse en un segundo idioma [40], [26], [23]
- Capacidad para comunicarse de forma oral y por medio de la palabra escrita en la lengua materna [28], [34], [20], [23], [10]
- Capacidad para generar nuevas ideas (creatividad) [34], [20], [10]
- Capacidad para trabajar en equipo [25], [28], [36], [21], [20], [23], [10]
- Capacidad para tomar decisiones razonadas [1], [10]
- Capacidad para planificar y administrar el tiempo [1], [20]
- Capacidad para adaptarse y actuar en nuevas situaciones [20]
- La apreciación y el respeto por la diversidad y multiculturalidad [20], [10]
- Capacidad de aplicar los conocimientos en situaciones prácticas [20]
- Capacidad para identificar, plantear y resolver problemas [1], [15], [10]
- Habilidad para trabajar de forma autónoma [10]
- Capacidad para el pensamiento abstracto, análisis y síntesis [10]
- Capacidad para ser críticos y autocríticos [10]
- Espíritu de empresa, la capacidad de tomar la iniciativa [10]

#### **4.1.2. ¿Qué herramientas o metodologías se utilizan para evaluar competencias mediante el uso de los entornos virtuales?**

Las herramientas o metodologías que se han empleado en nuestra selección primaria para cada competencia son las siguientes:

- Capacidad para comunicarse en un segundo idioma:
  - Facebook [40]
  - Second Life [26]
  - Grabación en vídeo [23]
- Capacidad para comunicarse de forma oral y por medio de la palabra escrita en la lengua materna:
  - Grabación en vídeo [23]
- Capacidad para trabajar en equipo:
  - Entorno de seguimiento de trabajo en equipo. JAMTART (Joe And Marks Tracking And Reporting Tool) [25]
  - Grabación en vídeo [23], [22]
  - Wikis [34], [21]
- Capacidad para tomar decisiones razonadas:
  - Tricuspoid [1], se trata de una plataforma digital diseñada específicamente para la autoevaluación de competencias empresariales.
- Capacidad para planificar y administrar el tiempo
  - Tricuspoid [1].
- Capacidad para generar nuevas ideas (creatividad):
  - Wikis, blogs [34]
- Capacidad para identificar, plantear y resolver problemas:
  - Tricuspoid [1].
  - Juego serio [15]

#### **4.1.3. ¿Con qué técnicas se pueden obtener evidencias (numéricas) objetivas del desarrollo de las siguientes competencias en un entorno virtual?**

- Capacidad para comunicarse en un segundo idioma:
  - Evaluación entre iguales y/o autoevaluación [40], [23]
  - Evaluación conjunta de varios profesores [26]
- Capacidad para comunicarse de forma oral y por medio de la palabra escrita en la lengua materna
  - Autoevaluación [20]
  - Evaluación entre iguales y/o autoevaluación [23]
  - Rúbricas [28]

- Capacidad para generar nuevas ideas (creatividad)
  - Autoevaluación [20]
- Capacidad para trabajar en equipo
  - Autoevaluación y/o evaluación entre iguales [25], [21], [23], [20]
  - Rúbricas [28], [34]
  - Técnica estadística: Modelo ESPEGS [36]
- Capacidad para tomar decisiones razonadas:
  - Autoevaluación y/o evaluación entre iguales [1]
- Capacidad para planificar y administrar el tiempo
  - Autoevaluación y/o evaluación entre iguales [1], [20]
- Capacidad para adaptarse y actuar en nuevas situaciones
  - Autoevaluación [20]
- La apreciación y el respeto por la diversidad y multiculturalidad:
  - Autoevaluación [20]
- Capacidad de aplicar los conocimientos en situaciones prácticas:
  - Autoevaluación [20]
- Capacidad para identificar, plantear y resolver problemas:
  - Autoevaluación y/o evaluación entre iguales [1]
  - Indicadores basados en rúbrica [15]
- Habilidad para trabajar de forma autónoma:
  - Autoevaluación y/o evaluación entre iguales [10]
- Habilidad para trabajar de forma autónoma:
  - Autoevaluación y/o evaluación entre iguales [10]
- Capacidad para el pensamiento abstracto, análisis y síntesis:
  - Autoevaluación y/o evaluación entre iguales [10]
- Capacidad para ser críticos y autocríticos:
  - Autoevaluación y/o evaluación entre iguales [10]
- Espíritu de empresa, la capacidad de tomar la iniciativa:
  - Autoevaluación y/o evaluación entre iguales [10]

#### **4.1.4. ¿Los resultados que se obtendrían de la evaluación de competencias genéricas de los alumnos mediante el uso de los registros de actividad de los entornos virtuales son indicadores fiables del desempeño de dichas competencias?**

No se han encontrado trabajos que aborden esta problemática de manera directa, aunque sí se han encontrado evidencias en la literatura que puedan ayudar a responder a esta pregunta. En alguno de los trabajos se utilizan wikis para fomentar el desempeño de los alumnos en diferentes competencias [34]. Lim indica que los tutores pueden recoger una gran cantidad de información acerca de sus estudiantes mediante la observación de desempeño de los alumnos, mediante la construcción de los wikis, los comentarios realizados por los estudiantes, y los intercambios entre estudiante [21]. Pero es evidente que si el curso tiene un gran número de alumnos la observación se vuelve inabordable. Nuestro trabajo va en la línea de observar el trabajo de los alumnos en los LMS mediante la extracción automática de indicadores. Aunque un LMS es más completo en cuanto al ámbito de actuación que un wiki, en ambos los usuarios trabajan en el sistema de forma independiente, dejando un rastro de su actividad y en ambos se sustenta el trabajo colaborativo. Es más, los LMS suelen incluir un wiki en su estructura. *Un wiki es un tipo de página web que brinda la posibilidad de que multitud de usuarios puedan editar sus contenidos a través del navegador web, con ciertas restricciones mínimas. De esta forma permite que múltiples autores puedan crear, modificar o eliminar los contenidos. Se puede identificar a cada usuario que realiza un cambio y recuperar los contenidos modificados, volviendo a un estado anterior. Estas características facilitan el trabajo en colaboración así como la coordinación de acciones e intercambio de información sin necesidad de estar presentes físicamente ni conectados de forma simultánea* (Wikipedia).

Por último, Cardona dice que la evaluación de las competencias implica la identificación de los elementos en torno a los procesos de aprendizaje, lo que hace que sea una actividad constante que requiere de criterios para evaluar los resultados durante la formación de las personas [8]. Consideramos que el rastro que dejan los alumnos, su interacción con el entorno, es un reflejo de su trabajo. Chebil indica como 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 [9]. Mientras que Florian argumenta que para poder utilizar esta información almacenada acerca de las actividades del curso, se requiere filtrado antes de que pueda ser utilizado para procesarla [12].

## Capítulo 5

# Trabajo (presente y) futuro

### 5.1. EvalCourse

En este trabajo hemos buscado literatura que aborde la evaluación de competencias genéricas, y a ser posible, que lo hagan de manera automática utilizando la interacción de los estudiantes con el LMS. En línea con este trabajo, y a fin de abordar lo que será mi futura tesis doctoral, estamos trabajando en una herramienta que cumpla este propósito. Esta herramienta, cuyo prototipo ya ha sido presentado en varios congresos [3, 4], se llama EvalCourse. EvalCourse es un Lenguaje de Dominio Específico (DSL). Un DSL es un lenguaje de programación orientado a un problema específico y con una semántica orientada al dominio para el que se diseña. En nuestro caso, este dominio es la evaluación de indicadores de competencias de los estudiantes. El objetivo de EvalCourse es ayudar al docente en la evaluación del desempeño de los alumnos en las competencias que éstos deben desarrollar a lo largo del curso. Estos datos deberán servir al docente como indicadores del desarrollo de dichas competencias. La idea inicial con la que se ha desarrollado este proyecto, es que el DSL sea genérico y pueda ser utilizado con otros sistemas, no sólo de tipo LMS, sino también con sistemas que fomenten el trabajo colaborativo, como pueden ser los wikis. De hecho, MediaWiki es utilizado también en el ámbito de la docencia de la Universidad de Cádiz.

### 5.2. Proyectos en curso

Para el presente curso se solicitó un proyecto en la convocatoria de actuaciones avaladas para la mejora docente, formación del profesorado y difusión de resultados, con el título *Extracción de indicadores objetivos para evaluación del desarrollo de competencias genéricas a partir de registros de actividad del Campus Virtual*. La Unidad de Innovación Docente de la Universidad de Cádiz nos concedió dicha actuación y actualmente estamos trabajando en este proyecto.

Este proyecto se engloba dentro del trabajo realizado dentro del grupo de investigación SPI-FM (TIC-195) por parte de Antonio Balderas al uso de las tecnologías en los entornos de aprendizaje. Se pretende desarrollar y aplicar EvalCourse, un DSL que permita a los docentes mediante el uso de una sintaxis sencilla obtener indicadores del desempeño de los estudiantes en diferentes competencias. Son varias las asignaturas de Grado en Ingeniería Informática que participan en este proyecto. Las tareas que comprende son:

1. Analizar los tipos de indicadores que se podrían obtener de la base de datos Moodle que sean de utilidad para los profesores, así como la manera de obtenerlos.
2. Análisis, desarrollo y prueba del software. Implementar el software DSL para que se conecte a Moodle y pueda extraer la información. Para ello sería necesario la instalación de un entorno

de pruebas Moodle y la generación de una batería de datos de pruebas que permita verificar el funcionamiento del software.

3. Obtención de datos reales de las asignaturas del campus virtual. A partir de estos, utilizaremos la herramienta para obtener todos los indicadores necesarios para evaluar las competencias de los estudiantes. Además podremos corregir aquellos fallos que se detecten en el software.
4. Análisis de resultados de la experiencia, difusión del trabajo realizado, publicación de informes, correcciones menores en el sistema, etc. Se analizarán los resultados obtenidos y se darán a conocer el software y la experiencia en webs especializadas, redes sociales, etc. Igualmente se procederá a la correcta documentación del sistema en una web oficial para poder replicar la experiencia (incidiendo en cómo evitar los errores cometidos), así como pequeños desarrollos que corrijan deficiencias observadas y faciliten la reutilización del sistema.

Los objetivos del proyecto son los siguientes:

1. Realizar un análisis en profundidad de qué indicadores se pueden extraer del campus virtual basado en Moodle que realmente reflejen el desarrollo o no de una competencia por parte de los alumnos.
2. Desarrollar un software basado en un DSL que permita el profesor de manera sencilla y escalable una evaluación objetiva del desarrollo de las competencias por parte de sus alumnos.
3. Aplicar la herramientas los datos extraídos de los cursos del campus virtual de las asignaturas participantes en el proyecto.
4. Análisis de los resultados obtenidos y difusión del trabajo realizado.

## Capítulo 6

# Conclusiones

En el presente documento se ha realizado una revisión formal de la literatura para responder a las preguntas de la investigación relacionadas con la evaluación automatizada de competencias genéricas mediante el uso de entornos de aprendizaje virtuales.

Los criterios de inclusión de artículos apenas nos dejaron un 7 % de los que fueron preseleccionados tras el proceso de búsqueda llevado a cabo en las diferentes bibliotecas digitales. La selección primaria ha sido clasificada en base al tipo de contribución que realizan a la ciencia, en base al tipo de investigación y en base al ámbito de aplicación del trabajo. Esta clasificación ha sido representada mediante diferentes tablas y figuras.

El escaso número de resultados obtenido viene a corroborar que es un tema de reciente interés, siendo además destacable que el grueso de artículos rescatado data de los últimos 3 años. No obstante, no se ha encontrado un trabajo de las características del que tenemos intención de realizar. Las competencias genéricas se evalúan mediante el uso de herramientas de tipo LMS, Wiki o Redes Sociales, pero casi siempre es necesario el factor humano para evaluar o estimar los indicadores de competencias. Normalmente mediante el uso de una rúbrica o mediante la revisión de una consecución satisfactoria o no de objetivos del curso.

Las preguntas de investigación planteadas fueron respondidas encontrando trabajos asociados a cada una de ellas. Aunque como se indica en el párrafo anterior, los trabajos encontrados evalúan competencias genéricas, aunque no lo hacen desde una perspectiva tan automatizada como se pretende que sea nuestro trabajo.

Se presentan las líneas de trabajo futuro, basadas en el desarrollo de la herramienta EvalCourse para la evaluación automática de indicadores de los LMS. En la actualidad hay un proyecto de Innovación Docente coordinado por el autor de este trabajo y que junto con este trabajo forman las bases de la que será su tesis doctoral.





# Bibliografía

- [1] Fariza Achcaoucaou, Laura Guitart-Tarrés, Paloma Miravittles-Matamoros, Ana Núñez Carballosa, Mercé Bernardo, and Andrea Bikfalvi. Competence assessment in higher education: A dynamic approach. *Human Factors and Ergonomics in Manufacturing & Service Industries*, pages n/a–n/a, June 2012.
- [2] Lorin W. Anderson, David R. Krathwohl, Peter W. Airasian, Kathleen A. Cruikshank, Richard E. Mayer, Paul R. Pintrich, James Rath, and Merlin C. Wittrock. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Abridged Edition, 2001.
- [3] A. Balderas, I. Ruiz-Rube, J. M. Dodero, M. Palomo-Duarte, and A. Berns. *A generative computer language to customize online learning assessments*, volume 8095 LNCS of *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer Berlin Heidelberg, 2013.
- [4] Antonio Balderas, Iván Ruiz-Rube, Manuel Palomo-Duarte, and Juan M. Dodero. A generative computer language to customize online learning assessments. In *Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality*, TEEM '13, pages 141–147, New York, NY, USA, 2013. ACM.
- [5] Michael Bedek, Sobah Abbas Petersen, and Tuija Heikura. From Behavioral Indicators to Contextualized Competence Assessment. In *2011 IEEE 11th International Conference on Advanced Learning Technologies*, pages 277–281. IEEE, July 2011.
- [6] Nancy Law Benő Csapó, John Ainley, Randy E. Bennett, Thibaud Latour. Technological Issues for Computer-Based Assessment. In Patrick Griffin, Barry McGaw, and Esther Care, editors, *Assessment and Teaching of 21st Century Skills*, pages 143–230. Springer Netherlands, 2012.
- [7] Fernando Borrajo, Yolanda Bueno, Isidro de Pablo, Begoña Santos, Fernando Fernández, Javier García, and Ismael Sagredo. SIMBA: A simulator for business education and research. *Decision Support Systems*, 48(3):498–506, February 2010.
- [8] Sergio Cardona, Jeimy Velez, and Sergio Tobon. Towards a model for assessing competencies. In *2013 8th Computing Colombian Conference (8CCC)*, pages 1–6. IEEE, August 2013.
- [9] 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.
- [10] Ricardo Colomo-Palacios, Cristina Casado-Lumbreras, Pedro Soto-Acosta, Francisco J. García-Peñalvo, and Edmundo Tovar-Caro. Competence gaps in software personnel: A multi-organizational study. *Computers in Human Behavior*, 29(2):456–461, March 2013.

- [11] Robert Feldt, Martin Höst, and Frank Lüders. Generic skills in software engineering master thesis projects: Towards rubric-based evaluation. In *2009 22nd Conference on Software Engineering Education and Training*, pages 12–15. IEEE, 2009.
- [12] 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.
- [13] Consolacion Gil, Maria G. Montoya, Rosario I. Herrada, Raul Banos, Francisco G. Montoya, and Francisco Manzano Agugliaro. Cooperative learning and electronic group portfolio: tutoring tools, development of competences and assessment. *International Journal of Learning Technology*, 6(1):46, May 2011.
- [14] Julia Gonzalez and Wagenaar Robert. *Tuning Educational Structures in Europe - Final Report - Pilot Project - Phase I*. University of Deusto, 2003.
- [15] M. Guenaga, S. Arranz, I. Rubio, E. Aguilar, A. Ortiz de Guinea, A. Rayon, M. Bezanilla, and I. Menchaca. Serious Games for the Development of Employment Oriented Competences. *Revista Iberoamericana de Tecnologías del Aprendizaje*, 2013.
- [16] 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.
- [17] 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.
- [18] D. Kennedy. *Writing and Using Learning Outcomes: A Practical Guide*. University College Cork, 2007.
- [19] 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.
- [20] C.H. Liao, M.H. Yang, and B.C. Yang. Developing a diagnosis system of work-related capabilities for students: A computer-assisted assessment. *Journal of Computer Assisted Learning*, pages n/a–n/a, April 2013.
- [21] Cher Ping Lim, Eugenia Ng, Yiu Chi Lai, and Eugenia M.W. Ng. Using wikis to develop student teachers’ learning, teaching, and assessment capabilities. *The Internet and Higher Education*, 14(1):15–26, 2011.
- [22] A.M. Martin-Cuadrado, M.A. Lopez-Gonzalez, and A. Garcia-Arce. Innovation Network: Video-conferencing as a Resource in Teaching Support and Autonomous Learning, 2013.
- [23] Albert Masip-Alvarez, Carme Hervada-Sala, Teresa Pamies-Gomez, Antoni Arias-Pujol, Carles Jaen-Fernandez, Cristina Rodriguez-Sorigue, David Romero-Duran, Fatiha Nejjari-Akhi-Elarab, Maria-Dolores Alvarez-del Castillo, Marta Roca-Lefler, Jorge Hernandez-Farras, and Raul Montferrer-Linan. Self-video recording for the integration and assessment of generic competencies. In *2013 IEEE Global Engineering Education Conference (EDUCON)*, pages 436–441. IEEE, March 2013.
- [24] C. McLoughlin and J. Luca. Beyond marks and measurement: Developing dynamic and authentic forms of e-assessment. In *Australian Society for Computers in Learning in Tertiary Education*, 2006.

- [25] Mark McMahon and Joe Luca. Explorations in metacognition: The design, development, and implementation of an online teamwork tracking environment, 2007.
- [26] J. Enrique Agudo Mercedes Rico, Julian Coppens, Paula Ferreira, Héctor Sánchez. Everything Matters: Development of Cross-Curricular Competences in Engineering Through Web 2.0 Social Objects. In Demetrios G. Sampson, Pedro Isaias, Dirk Ifenthaler, and J. Michael Spector, editors, *Ubiquitous and Mobile Learning in the Digital Age*, pages 139–157. Springer New York, 2013.
- [27] Azlinah Mohamed, Azrilah Aziz, Sohaimi Zakaria, and Mohd Saidfudin Masodi. Appraisal of course learning outcomes using rasch measurement: a case study in information technology education. In *SEPADS'08 Proceedings of the 7th WSEAS International Conference on Software Engineering, Parallel and Distributed Systems*, volume 1, pages 232–238. World Scientific and Engineering Academy and Society (WSEAS), 2008.
- [28] Azlinah Mohamed, Azrilah Abd Aziz, Abd Rahman Ahlan, Sohaimi Zakaria, and Mohd Saidfudin Masodi. Outcome based education performance measurement: a rasch-based longitudinal assessment model to measure information management courses lo's. *WSEAS Transactions on Information Science and Applications*, 5(3):292–299, 2008.
- [29] 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.
- [30] Deborah Murdoch-Eaton and Sue Whittle. Generic skills in medical education: developing the tools for successful lifelong learning. *Medical education*, 46(1):120–8, January 2012.
- [31] Acensión Palomares. The educational model at university and the use of new methodologies for teaching, learning and assessment. *Revista de Educación*, 355(Mayo-Agosto 2011):591–604, 2011.
- [32] 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.
- [33] Sobah A. Petersen and Michael A. Bedek. Challenges and Opportunities in Evaluating Learning in Serious Games: A Look at Behavioural Aspects. In *Serious Games Development and Applications*, pages 219–230. Springer Berlin Heidelberg, January 2012.
- [34] Nelson Piedra, Janneth Chicaiza, Jorge Lopez, Audrey Romero, and Edmundo Tovar. Measuring collaboration and creativity skills through rubrics: Experience from UTPL collaborative social networks course. In *IEEE EDUCON 2010 Conference*, pages 1511–1516. IEEE, 2010.
- [35] Cachia R., Ferrari A., Ala-Mutka K., and Punie Y. *Creative Learning and Innovative Teaching: Final Report on the Study on Creativity and Innovation in Education in the EU Member States*. Dictus Publishing, 2011.
- [36] Rozeha A. Rashid, Razimah Abdullah, Azami Zaharim, Hamzah Ahmad Ghulman, and Mohd Saidfudin Masodi. Engineering students performance evaluation of generic skills measurement: Espes model. In *5th WSEAS / IASME International Conference on ENGINEERING EDUCATION (EE'08)*, pages 377–383. World Scientific and Engineering Academy and Society (WSEAS), 2008.
- [37] Christine Redecker and Øystein Johannessen. Changing assessment — towards a new assessment paradigm using ict. *European Journal of Education*, 48(1):79–96, 2013.

- [38] Christine Redecker, Yves Punie, and Anusca Ferrari. eAssessment for 21st Century Learning and Skills. In Andrew Ravenscroft, Stefanie Lindstaedt, Carlos Delgado Kloos, and Davinia Hernández-Leo, editors, *21st Century Learning for 21st Century Skills*, volume 7563 of *Lecture Notes in Computer Science*, pages 292–305. Springer Berlin Heidelberg, September 2012.
- [39] D. Sampson and D. Fytros. Competence models in technology-enhanced competence-based learning. In HeimoH. Adelsberger, Kinshuk, JanM. Pawlowski, and DemetriosG. Sampson, editors, *Handbook on Information Technologies for Education and Training*, International Handbooks on Information Systems, pages 155–177. Springer Berlin Heidelberg, 2008.
- [40] Ru-Chu Shih. Can Web 2.0 technology assist college students in learning English writing? Integrating Facebook and peer assessment with blended learning, 2011.
- [41] Beth Simon and Jared Taylor. What is the value of course-specific learning goals? *College Science Teaching*, 39(2):52–57, 2009.
- [42] Andreja Istenic Starcic. Developing virtual simulation game for authentic learning: realizing partnership between university and industry. *WSEAS TRANSACTIONS on COMMUNICATIONS*, 7(7):786–795, July 2008.
- [43] Andreja Istenič Starčič. Game-based learning in higher education and lifelong learning: bridging the gap between theory and practice. *Game-based learning in higher education and lifelong learning: bridging the gap between theory and practice*, pages 157–162, September 2008.
- [44] MaríaJosé Terrón-López and MaríaJosé García-García. Assessing transferable generic skills in language degrees. In *Competency-based Language Teaching in Higher Education*, volume 14 of *Educational Linguistics*, pages 151–167. Springer Netherlands, 2013.
- [45] Manuel Salas Velasco, María Teresa Sánchez Martínez, and Noelina Rodríguez Ferrero. Developing Generic Competences in the European Higher Education Area: a proposal for teaching the principles of economics. *European Journal of Education*, 47(3):462–476, September 2012.
- [46] 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.
- [47] 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.
- [48] 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.