

Evolución de los MOOC en el ámbito investigador mediante técnicas de análisis de contenido

D. G. Reina, S. L. Toral
Dpto. Ingeniería Electrónica
University of Seville
Sevilla, España
{dgutierrezreina,storl}@us.es

M. R. Martínez-Torres, F. Barrero
Dpto. Ingeniería Electrónica
University of Seville
Sevilla, España
{rmtorres,fbarrero}@us.es

Abstract— Los Massive Open Online Courses (MOOCs) constituyen un nuevo modelo de docencia que ha ido ganando popularidad en los últimos años. Su carácter masivo implica un modelo de enseñanza colaborativa que genera conexiones no sólo entre el profesor y el alumno, sino también entre los propios alumnos, que intervienen en foros y participan en procesos de evaluación por pares. El ámbito de la investigación en educación no es ajeno a este auge, y el campo de los MOOC también es un tema emergente, que ofrece además múltiples retos investigadores en cuanto a su modelo de desarrollo, su impacto en los modelos tradicionales de educación superior o su gestión dado su carácter masivo. Aunque en los últimos años son varias las revistas que han lanzado varios números monográficos sobre el tema de los MOOCs, no existe un estudio bibliográfico hasta la fecha que ilustre sus principales retos investigadores. Este artículo propone la realización de un estudio bibliográfico basado en técnicas de análisis de contenido que muestre no sólo los principales ámbitos de interés sino también su evolución en el tiempo.

Keywords— MOOC; Planes de estudio; Análisis de contenidos; Análisis cluster

I. INTRODUCCIÓN

Los MOOCs son cursos abiertos que se imparten de manera online con un contenido públicamente compartido y accesible por los alumnos que se matriculan [1]. Su éxito radica en su capacidad de atraer miles de estudiantes de todo el mundo que se organizan su propio aprendizaje y establecen conexiones no sólo con los docentes sino con otros estudiantes que participan en el curso. En general los MOOCs se basan en las mismas premisas que ha popularizado la enseñanza online en el sentido de eliminar barreras espacio-temporales y facilitar el uso de recursos de manera virtual. No obstante, en la actualidad representan un campo diferenciado dentro de los recursos educativos abiertos, y desde hace años se han comenzado a estudiar separadamente en el ámbito de la investigación [2]. Entre otros aspectos se han analizado los motivos de su éxito, el perfil de los alumnos que se matriculan, los retos tecnológicos relacionados con su carácter masivo, el bajo ratio de alumnos que lo finalizan, los procesos pedagógicos que se desarrollan o los sistemas de evaluación dado su carácter masivo. Su popularidad ha hecho que muchas universidades e instituciones de educación superior vuelvan sus ojos hacia esta novedosa posibilidad y comienzan a incluirla dentro de su

oferta formativa. Al ser un tema multidisciplinar y que abarca multitud de aspectos, esta investigación es todavía dispersa, por lo que es necesario un estudio bibliográfico que aclare a futuros investigadores los principales campos de interés y su evolución en el tiempo. Con ese fin, este estudio propone la aplicación de técnicas de análisis de contenido para determinar cuáles son los principales temas de interés en el ámbito de los MOOCs durante los últimos seis años (2010-2015).

II. CARACTERÍSTICAS DE LOS MOOCs

Los MOOCs están ganando atención en los últimos años gracias al elevado número de estudiantes que participan y sus alianzas con universidades prestigiosas que ofrecen cursos gratuitos o a muy bajo coste a potenciales estudiantes por todo el mundo [3]. El rápido auge de los MOOC ha planteado varias cuestiones de investigación, muchas de ellas relacionadas con su desarrollo y su encaje dentro del modelo tradicional de enseñanza de la educación superior. A diferencia de la enseñanza reglada, los cursos ofertados presentan una gran variedad y son flexibles y adaptables rápidamente a las temáticas más novedosas y demandadas [4]. Además, son cursos dirigidos a una audiencia a escala mundial. No importa lo específico o multidisciplinar que sea el curso, la economía de escala que proporciona una potencial audiencia de miles de estudiantes posibilita el efecto de cola larga, de modo que cursos que a nivel local no conseguirían un mínimo número de estudiantes pueden ser viables en forma de MOOC [5].

Uno de los aspectos más estudiados en el ámbito de los MOOCs es el perfil de los estudiantes que los cursan. Por lo general suelen ser estudiantes ya titulados o de estudios avanzados que quieren mejorar su formación en temas concretos. Según estudios recientes, dos terceras partes de los alumnos se identifican como empleados, y un 44% los cursan para poder hacer mejor su trabajo [6]. No obstante y a pesar del alto número de estudiantes que se apuntan a estos cursos, también es cierto que una gran mayoría no llega a completarlos, por falta de tiempo, incentivos, conocimiento previo insuficiente.

Los MOOCs también prestan especial atención a las interacciones entre los usuarios, normalmente a través de foros. A través de ellos, los usuarios comentan temas, ejercicios, proponen nuevos tópicos de discusión o incluso interactúan con los instructores para resolver algunas cuestiones relacionadas

con la asignación de trabajos. Estudios previos en esta cuestión han concluido que la interacción de los usuarios a través de los foros mejora la comprensión de los contenidos [8].

Los sistemas de evaluación empleados también forman parte de los temas de investigación, teniendo en cuenta que se trata de cursos masivos donde los sistemas de evaluación tradicionales requerirían de un tiempo prohibitivo. Habitualmente las evaluaciones son, o bien automáticas, mediante una herramienta software que compruebe las respuestas a un test o evalúe el resultado de un programa, o mediante un sistema de evaluación por pares donde son los mismos estudiantes matriculados en el curso los que evalúan los trabajos enviados por otros estudiantes, siguiendo unas determinadas pautas previamente indicadas en el curso. La ventaja de la evaluación por pares es que permite evaluar trabajos abiertos, donde los alumnos pueden desarrollar más su creatividad [9], aunque a costa de pedirles una mayor carga de trabajo. Actualmente, existe un gran debate acerca de los sistemas de evaluación automatizada, AES (Automated Essay Scoring) [10].

Como todos los temas emergentes, la investigación en el ámbito de los MOOCs se encuentra en un estadio temprano, lo que provoca que sea aún dispersa. En este sentido, son importantes los estudios que tratan de estructurar los principales temas de investigación tratados hasta la fecha, con el fin de ayudar a los investigadores actuales y a los futuros a elegir los temas más adecuados en el contexto de lo realizado hasta la fecha. En esta línea cabe destacar el trabajo realizado en [11], donde se lleva a cabo un análisis de los principales temas de interés en el ámbito de los MOOCs a partir de los textos de los proyectos de investigación presentados dentro de la convocatoria del MOOC Research Initiative. En este caso, se utilizará como cuerpo documental las publicaciones científicas recogidas en el ISI Web of Knowledge que recojan los términos MOOCs o Massive Open Online Courses dentro de sus tópicos de interés, título o resumen

III. METODOLOGÍA

El análisis de contenido consiste en la aplicación de métodos computacionales, basados en el procesamiento del lenguaje natural y minería de textos, para obtener los principales tópicos de interés dentro de un cuerpo documental. Normalmente estas técnicas se utilizan cuando se trabaja con un cuerpo documental muy grande, donde es difícil extraer estos temas clave de una forma manual. La aplicación de técnicas computacionales también evita sesgos introducidos por las personas que analizan manualmente los textos. De entre las diferentes técnicas de análisis de contenido, la más utilizada es la que se basa en la co-ocurrencia de términos clave en los diferentes documentos que forman el cuerpo documental de análisis [12]. El primer paso para aplicar esta metodología consiste en capturar un conjunto de términos relevantes sobre los que basar la identificación de los tópicos tratados en el cuerpo documental. Para esta captura se utiliza la métrica conocida como TF*IDF (Term frequency–Inverse document frequency), que es ampliamente utilizada para evaluar la relevancia de términos claves en una colección de documentos [13]. El valor de TF considera que la importancia de un término es proporcional a su frecuencia en una colección de

documentos, en tanto que el valor IDF considera que la importancia de un término relativa a un documento es inversamente proporcional a la frecuencia de ocurrencia de un término en todos los documentos. La combinación de ambos valores define la importancia de un término según su frecuencia pero teniendo en cuenta que un término que aparece en una mayoría de documentos deja de ser representativo por trivial. A partir del valor TF-IDF se pueden extraer tanto términos clave como frases clave de un número variable de palabras.

Una vez obtenidos los términos o frases clave, se puede llevar a cabo el análisis de tópicos. Para ello se elabora en primer lugar una matriz de frecuencias términos-documentos sobre la que se aplica un análisis factorial con rotación Varimax. A diferencia del análisis clúster jerárquico, donde una palabra sólo puede aparecer asociada a un tópico, el análisis de tópicos basado en el análisis factorial permite que una palabra se pueda asociar con más de un factor. Esta característica es importante, ya que permite considerar la naturaleza polisémica de muchas palabras y la variedad de contextos en el que un término se puede utilizar [14].

Finalmente, se utilizará un análisis de correspondencia para visualizar la relación entre los términos y frases clave extraídas con los años 2010-2015, con el fin de analizar su evolución temporal [15].

IV. RESULTADOS

Para el análisis de la evolución de los MOOCs en el ámbito científico se ha utilizado como fuente las publicaciones científicas recogidas en el ISI Web of Knowledge que incluyen los términos MOOCs o Massive Open Online Courses dentro de sus tópicos de interés, título o resumen para el periodo comprendido entre 2010 y 2015.

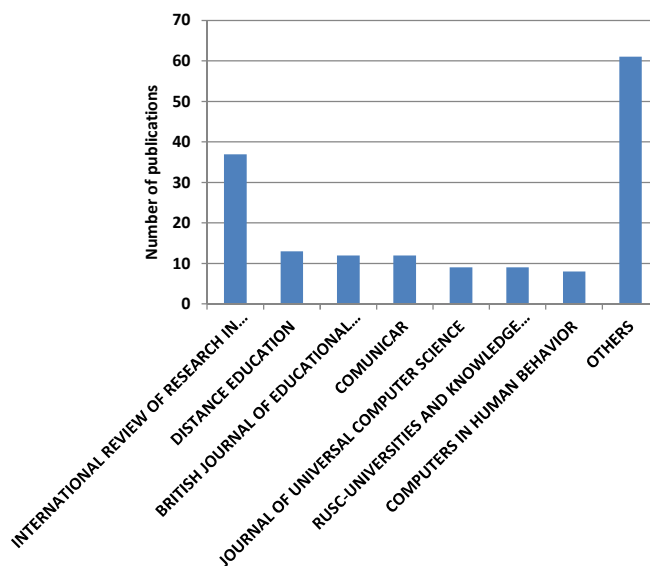


Figura 1. Distribución de publicaciones relacionadas con los MOOCs en revistas internacionales.

Se han considerado tanto revistas como congresos educativos internacionales.

La Figura 1 muestra como las publicaciones obtenidas se reparten entre las diferentes revistas indexadas. Las publicaciones se reparten entre un total de 24 revistas diferentes, y como se desprende de la Figura 1, se trata de un tópico bastante transversal. Las 8 primeras revistas acumulan un 62% de las publicaciones, y el 38 % restante está muy repartido. Las dos revistas que más publicaciones acumula son el *International Review of Research in Open and Distance Learning* y el *Distance Education*. En ambos casos, la educación abierta y a distancia es uno de sus tópicos principales.

La selección de términos clave se llevó a cabo considerando los 100 términos más relevantes según el TF-IDF y eliminando aquellos términos que aparecen en más del 50% de los documentos. De este modo se consigue eliminar términos triviales tales como MOOC, Massive o Open. El análisis de tópicos se llevó a cabo utilizando el análisis factorial propuesto con una rotación Varimax. La Tabla I muestra el resultado obtenido para un total de 9 clústeres, detallando las palabras clave que incluye cada clúster y el porcentaje de varianza explicado. Como principales tópicos en el ámbito de la investigación se pueden destacar los siguientes:

- El primero de los clústeres se centra en los restos y oportunidades que ofrecen los MOOCs en el ámbito educativo. Por ejemplo, los MOOCs poseen una tasa de matriculación muy alta pero muy baja de eficiencia, dado que la mayoría no llega a finalizarlos y no participan en las distintas actividades que se organizan. Esto se debe a que requieren un esfuerzo y tiempo por parte de los estudiantes y que ellos mismo se regulan su aprendizaje. Por otro lado los términos “*research*” y “*paper*” hacen referencia al carácter investigador de los MOOC como herramientas educativas de nueva generación.
- Otra línea diferente de trabajo se refiere a los contenidos que pueden ofertarse dentro de los MOOCs. Típicamente, el material consiste en videos y las transparencias en pdf. No obstante, existen también otras posibilidades en cursos más especializados o con contenidos de carácter práctico [16].
- Las interacciones que se producen entre estudiantes y entre estos y los instructores es uno de los puntos fuertes de los MOOCs. Estas interacciones se han estudiado frecuentemente en la literatura desde el punto de vista de las comunidades que pueden resultar de estas interacciones [17]. Los MOOCs hacen posible la interacción de estudiantes internacionales de todo el mundo con universidades de alto prestigio.
- El cuarto clúster se refiere a las diferentes metodologías de aprendizaje que pueden desarrollarse mediante los MOOCs. En general, son metodologías de aprendizaje a distancia, aunque en el caso de los MOOCs se combina también con un aprendizaje colaborativo través de los foros de discusión. Los estudiantes de un mismo MOOC y sus interacciones pueden ser percibidas como una red social de usuarios interesados en la misma temática.

- El clúster 5 se encuentra relacionado con el 3, pero se refiere a los trabajos que han estudiado las interacciones de los usuarios a través del análisis de redes sociales. En este sentido, se han analizado los usuarios que ejercen un rol central dentro de la red de usuarios o los patrones de comportamiento de usuarios e instructores [18].
- El clúster sexto se refiere a la experiencia de enseñanza desde la perspectiva de los instructores. En este sentido se ha analizado las diferencias entre la enseñanza reglada y la enseñanza ofertada en modalidad MOOC, así como las razones por las cuales los instructores deciden ofertar MOOCs [19].
- Los retos tecnológicos que afrontan los MOOCs en cuanto a que son sistemas masivos de aprendizaje también constituyen un tópico de estudio. Básicamente existen dos tendencias. La primera sigue una tendencia constructivista, individualista y basada en los sistemas de gestión del aprendizaje. La segunda se basa en el aprendizaje social, la cooperación y la web 2.0 [20].
- El clúster 9 abarca las actividades de evaluación dentro de los MOOCs. Dado el carácter masivo de los cursos, las evaluaciones se llevan a cabo mediante pares, usando otros estudiantes matriculados en el curso, o mediante herramientas automáticas de evaluación [21]. Para ello es esencial la colaboración y predisposición de los estudiantes del curso.
- Finalmente, el clúster 9 se refiere a la integración de los MOOCs dentro de la enseñanza reglada de los sistemas de Educación Superior. En este sentido, sistemas híbridos o incluso universidades online se contemplan como posibilidades de carácter más abierto [22].

Tabla I. Clústeres resultado de los análisis de tópicos sobre los términos clave.

	Name	Keywords	%Var
1	Challenges and opportunities	Challenge; Issue; Opportunity; Focus; Environment; Paper; Research; Practice	3,45
2	MOOC contents	Video; Lecture; Content; Present; Material; Paper; Experience; User; Platform; Create; Development; Support; Develop	2,91
3	Student interactions	Study; Participant; Result; Group; Interaction; Analysis; Discussion; Learner; Survey; Research; Provide; Design	3,62
4	Learning methodologies	Method; Base; Model; System; Data; Problem; Author; Show; Result; Approach; Analyze	2,77
5	Social networking	Social; Network; Medium; Impact	3,12
6	Teaching experience	Teach; Class; Traditional; University; Student; Teacher; Face; College; Group	2,76
7	Technological	Question; Project; Training;	2,08

	framework	Management; Time; Tool; Technology; Result	
8	Assessment	Peer; Assessment; Review; Number; Increase; Large; Student; Improve; Level; Cost	3,05
9	Integration in Higher Education	Public; Field; High; Year; Medium; Program; Science; Analysis; Impact; Analyze	2,62

Un análisis similar se ha realizado considerando diferentes conjuntos de palabras de tamaño mayor que uno. Evidentemente, este análisis sólo considera aquellas palabras que suelen aparecer juntas, y por tanto es más restrictivo que el análisis anterior, ya que puede ocurrir que términos clave muy frecuentes aparezcan en diferentes combinaciones de palabras, haciendo que no aparezcan en los resultados de este análisis. La Figura 2 muestra el resultado de este análisis en forma de dendrograma.

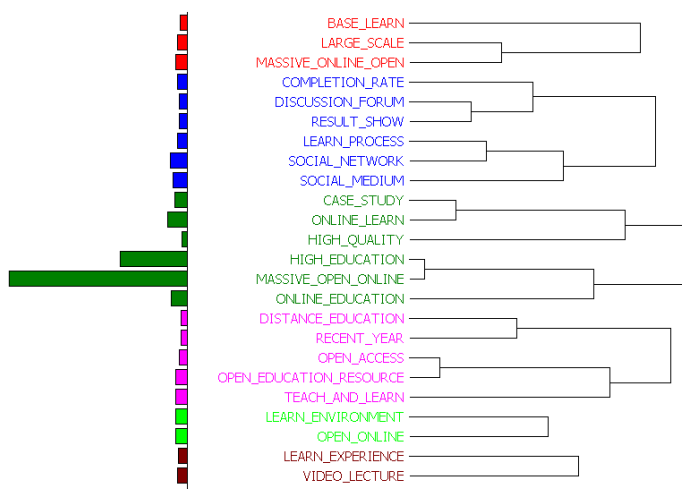


Figura 2. Dendrograma de las frases más relevantes según su valor TF-IDF.

De arriba abajo, se distinguen temas relacionados con los siguientes puntos:

- Carácter masivo de los cursos.
- Interacciones sociales.
- Integración con enseñanza de educación superior
- Procesos de enseñanza-aprendizaje en formato abierto
- Entorno tecnológico de aprendizaje
- Contenidos ofertados y experiencia de aprendizaje.

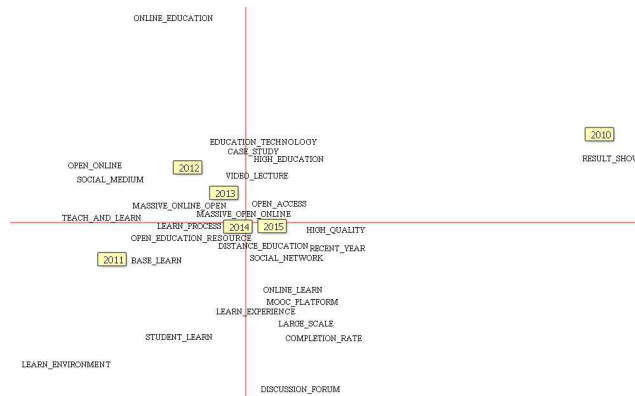


Figura 3. Mapa de correspondencia de las frases más relevantes durante el período 2010-2015.

La Figura 3 muestra el mapa de correspondencia que relaciona las frases más relevantes con los 6 años comprendidos entre 2010 y 2015 considerados en el estudio. Como suele ocurrir con los temas emergentes, los dos primeros años ocupan posiciones más periféricas, y abarcan los temas relacionados con resultados de los MOOCs y los procesos de enseñanza aprendizaje. A partir de 2012 comienzan a estudiarse las interacciones de los usuarios mediante análisis de redes sociales, contenidos, el formato abierto de los cursos y las plataformas tecnológicas de soporte. También es importante el carácter abierto “open access” de este tipo de cursos. Este término gana más importancia recientemente debido a que muchas plataformas como coursera están empezando a limitar el número de cursos totalmente abiertos. Se está pasando a un sistema híbrido, en cual algunos cursos son abiertos y otro no.

V. CONCLUSIONES

Este artículo ha analizado la evolución de los MOOCs en el ámbito investigador durante los últimos seis años utilizando técnicas de análisis de contenido. A partir de un conjunto de palabras y expresiones clave, se han obtenido los principales tópicos de investigación así como su evolución temporal. Los resultados obtenidos pueden servir de guía para futuros investigadores o para identificar terrenos donde todavía no se ha desarrollado la investigación relacionada con los MOOCs.

REFERENCES

- [1] T. Daradoumis, R. Bassi, F. Xhafa, & S. Caballe, “A Review on Massive E-Learning (MOOC) Design, Delivery and Assessment”, 2013 Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), Compiègne, Franc, 2013, pp. 208- 213
- [2] B. Toven-Lindsey, R. A. Rhoads, & J. Berdan Lozano, “Virtually unlimited classrooms: Pedagogical practices in massive open online courses”, The Internet and Higher Education, Vol. 24, pp. 1-12, 2015.
- [3] S. Kellogg, S. Booth, & K. Oliver, “A social network perspective on peer supported learning in MOOCs for educators”, The International Review of Research in Open and Distributed Learning, Vol. 15, Iss. 5, 2014.
- [4] S. Audsley, K. Fernando, B. Maxson, B. Robinson & K. Varney, "An Examination of Coursera as an Information Environment: Does Coursera Fulfill its Mission to Provide Open Education to All?", The Serials Librarian, Vol. 65, No. 2, pp. 136-166, 2013.

- [5] I. S. Abeywardena, "Public opinion on OER and MOOC: a sentiment analysis of twitter data", *Proceeding of International Conference on Open and Flexible Education (ICOFE 2014)*, Hong Kong SAR, China, 2014.
- [6] A. W. Radford, J. Robles, S. Cataylo, L. Horn, J. Thornton, & K. E. Whitfield, "The employer potential of MOOCs: A mixed-methods study of human resource professionals' thinking on MOOCs", *The International Review of Research in Open and Distributed Learning*, Vol. 15, Iss. 5, pp. 1-25, 2014.
- [7] K. F. Hew, W. S. Cheung, "Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges", *Educational Research Review*, Vol. 12, pp. 45-58, 2014.
- [8] D. K. Comer, C. R. Clark & D. A. Canelas, "Writing to learn and learning to write across the disciplines: Peer-to-peer writing in introductory-level MOOCs", *The International Review of Research in Open and Distributed Learning*, Vol. 15, Iss. 5, 2014.
- [9] N. J. Pelaez, "Calibrated peer review in general education undergraduate human physiology", In P. A. Rubba, J. A. Rye, W. J. DiBiase, & B. A. Crawford (Eds.), *Proceedings of the Annual International Conference of the Association for the Education of Teachers in Science*, Costa Mesa, CA, pp. 1518-1530, 2001.
- [10] S. P. Balfour, "Assessing Writing in MOOCs: Automated Essay Scoring and Calibrated Peer Review", *Research & Practice in Assessment*, Vol. 8, Iss.1, pp. 40-48, 2013.
- [11] D. Gasevic, V. Kovanovic, S. Joksimovic, & G. Siemens, "Where is research on massive open online courses headed? A data analysis of the MOOC Research Initiative", *The International Review Of Research In Open And Distributed Learning*, Vol. 15, Iss. 5, 2014.
- [12] B. Cheng, M. Wang, A. I. Mørch, N.-S. Chen, & J. M. Spector, "Research on e-learning in the workplace 2000–2012: A bibliometric analysis of the literature", *Educational Research Review*, Vol. 11, pp. 56–72, 2014.
- [13] N. Azam, J. T. Yao, "Comparison of term frequency and document frequency based feature selection metrics in text categorization," *Expert Systems with Applications*, vol. 39, no 5, pp. 4760-4768, 2012.
- [14] J.M. Gerken, M.G. Moehrl, "A new instrument for technology monitoring: novelty in patents measured by semantic patent analysis", *Scientometrics*, Vol. 91, pp. 645–670, 2012.
- [15] M. Greenacre, *Correspondence analysis in practice*, Vol. 2, New York: Chapman & Hall CRC, 2007.
- [16] D. Lowe, "MOOLs: Massive open online laboratories: An analysis of scale and feasibility", 2014 11th International Conference on Remote Engineering and Virtual Instrumentation (REV), IEEE, 2014. pp. 1-6.
- [17] V. Tinto, "Classrooms as communities: Exploring the educational character of student persistence", *The Journal of Higher Education*, 68(6), 599–623, 1997.
- [18] D. Gasevic, V. Kovanovic, S. Joksimovic & G. Siemens, "Where is research on massive open online courses headed? A data analysis of the MOOC Research Initiative", *The International Review Of Research In Open And Distributed Learning*, Vol. 15, no. 5, 2014.
- [19] K. F. Hew, & W. S. Cheung, Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges. *Educational Research Review*, Vol. 12, pp. 45-58, 2014.
- [20] Á. Fidalgo Blanco, M. L. Sein-Echaluce Laclleta, & F. J. García-Peñalvo, "Methodological Approach and Technological Framework to break the current limitations of MOOC model", *Journal of Universal Computer Science*, 21(5), 712-734, 2015.
- [21] E. D. Reilly, R. E. Stafford, K. M. Williams, & S. B. Corliss, "Evaluating the validity and applicability of automated essay scoring in two massive open online courses", *The International Review of Research in Open and Distributed Learning*, 15(5), 83-98, 2014.
- [22] TEL-Map (2012), TEL-Map UK HE Scenarios, <http://www.learningfrontiers.eu/?q=content/context-scenarios-task-7-2>