

How to map learning activities through URLs? The case of Coursera platform

Ronald Pérez-Álvarez^{1,2}, Jorge Maldonado-Mahauad^{1,3}, Mar Pérez-Sanagustín^{1,4}

¹ Pontificia Universidad Católica de Chile, Santiago, Chile

² Universidad de Costa Rica, Sede Regional del Pacífico, Puntarenas, Costa Rica

³ Universidad de Cuenca, Cuenca, Ecuador

⁴ Université Toulouse III Paul Sabatier, Toulouse, France

{raperez13, jjmaldonado, mar.perez}@uc.cl
mar.perez-sanagustin@irit.fr

Abstract. The increased use of online learning environments such as MOOCs creates an ideal setting for data collection and the application of learning analytics techniques to better understand how we learn. However, access to data from today's MOOC platforms is limited, and it is often difficult to collect and clean up data to obtain information that will help to better understand how learning occurs in these contexts. In this study we analyze how to track the learning traces of the students' activities on the MOOC Coursera platform to understand their behavior. The study analyzed the URL patterns of 13 MOOCs offered by the Pontificia Universidad Católica de Chile on this platform. As a result, 5 categories of activities were identified in the URLs: 1) the actions of the students on the platform; 2) the content involuted in the learning process; 3) the context in which the learning took place, and 4) the social interaction with other students. These results aim to provide guidance to other researchers in analyzing the behavior of students at MOOCs de Coursera and applying learning analytic techniques.

Keywords: Tools, learning analytics, Massive Open Online Courses, MOOC, Coursera, learning activities, MOOC platforms.

1 Introducción

El incremento en el uso de ambientes de aprendizaje online, como el caso de los MOOCs (Massive Open Online Courses), ha creado un escenario ideal para la recolección de datos y la incorporación de técnicas de analítica del aprendizaje [1]. Sin embargo, el acceso a los datos de las actividades de los estudiantes en la plataforma MOOC es limitado. Algunas plataformas como Coursera, edX, Future Learn proveen un paquete de datos con los logfiles de las actividades. No obstante, este paquete de datos se genera de forma periódica, con tiempos que varían entre un 1 a 15 días. Por ejemplo, el paquete de datos de Coursera se actualiza cada 24 horas, es decir las actividades de las últimas 24 horas no se encuentran en el paquete de datos. Kulkarni et al. [2] encuentran que los estudiantes que reciben retroalimentación en un rango menor a las 24 horas después de haber realizado sus actividades obtienen mejores

salidas de aprendizaje. Este desfase en los datos obtenidos de las plataformas dificulta el desarrollo de herramientas que entreguen retroalimentación inmediata y oportuna a los estudiantes, mientras realizan sus actividades de aprendizaje en la plataforma MOOC.

Bajo este panorama, se hace necesaria la incorporación de técnicas de recolección y seguimiento de las actividades de aprendizaje de los estudiantes en este contexto, con el fin de ofrecer retroalimentación en tiempo real a los estudiantes. Una alternativa para el seguimiento de las actividades de los estudiantes es mediante el análisis de los URLs visitados dentro de la plataforma MOOC, pero no todas las plataformas ofrecen suficiente información en los URLs para identificar que actividad realizó el estudiante. Actualmente, no hay investigaciones que nos permitan conocer cómo mapear la información obtenida en los URLs que visita el estudiante en la plataforma MOOC con las actividades de aprendizaje propuestas en el curso. Davis et al [3] diseñaron una herramienta llamada SRLx que sigue las actividades de los estudiantes en la plataforma edX y entrega retroalimentación en tiempo real, sin embargo, no se detalla información de cómo otros investigadores pueden mapear estas actividades.

En este estudio se realiza un análisis de los URLs que visita el estudiante en la plataforma MOOC Coursera y se identifican los componentes principales que permiten dar seguimiento las trazas de aprendizaje de las actividades de los estudiantes en la plataforma. Los resultados de esta investigación sirven como guía para futuras investigaciones dirigidas al análisis del comportamiento de los estudiantes en esta plataforma MOOC y así aplicar técnicas de analítica del aprendizaje para ofrecer retroalimentación en tiempo real a los estudiantes.

2 Cursos en la plataforma Coursera

La plataforma Coursera (<https://www.coursera.org>) fue lanzada en 2012 y actualmente cuenta con más de 30 millones de estudiantes registrados, alrededor de 150 universidades asociadas y más de 2.000 cursos disponibles [4] [5]. La Figura 1. muestra la estructura básica utilizada en el diseño de los cursos y la interfaz principal para la navegación del estudiante. A partir de esta página, el estudiante inicia su ruta de navegación entre las distintas actividades de aprendizaje. El menú principal derecho ofrece 4 opciones de navegación: **1)** mostrar la página de inicio del curso y el listado de semanas que componen el curso; **2)** revisar las calificaciones obtenidas en el curso; **3)** Acceder a los foros de debate y **4)** ver la información del curso.

El contenido de un curso se organiza en capítulos o módulos. Cada módulo es organizado en secciones o lecciones y estas a su vez se organizan en secuencias de aprendizaje. En Coursera cada semana del curso corresponde a un módulo (Figura 1), este curso se compone de 4 módulos, es decir tiene una duración de 4 semanas. Las secciones o lecciones de aprendizaje representan los temas a estudiar en cada semana. Una lección esta asociada a un módulo específico. Por último, una secuencia de aprendizaje es un conjunto de actividades a realizar que el profesor ha propuesto a los estudiantes como parte de una lección, tales como: evaluaciones, video lecciones, lecturas, tareas programadas, entre otras.

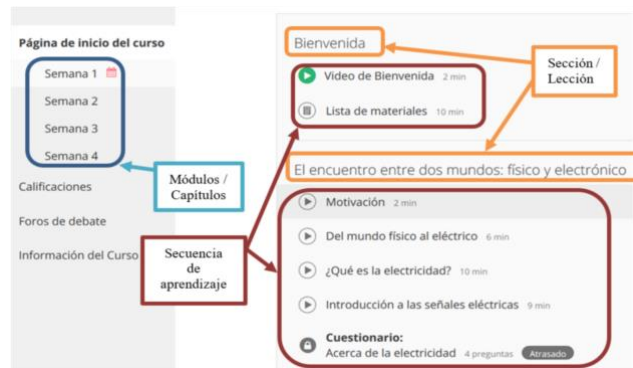


Fig. 1. Interfaz principal y organización de los cursos en Coursera.

Las actividades incluidas en las secuencias de aprendizaje son las actividades más relevantes para el seguimiento dentro de la plataforma, dado que estas corresponden propiamente a las actividades que realiza el estudiante durante su proceso de aprendizaje en el curso. El seguimiento de estas actividades permite obtener información relevante para aplicar técnicas de analítica del aprendizaje. Coursera utiliza 16 tipos de actividades de aprendizaje (Tabla 1) que el profesor puede agregar al diseñar una lección en la plataforma. Los foros de discusión no forman parte de este grupo de actividades, en Coursera se manejan como un tipo de actividad separada.

Tabla 1. Tipos de actividades para las secuencias de aprendizaje

id	Descripción	Categoría
1	Lecture	lecture
7	Others	--
10	graded programming	programming
8	graded peer	peer
9	closed peer	peer
14	ungraded lti	lti
12	phased peer	peer
11	ungraded programming	programming
15	Slideshow	slideshow
4	Peer	peer
2	assess open single page	quiz
16	staff graded	staffGraded
5	Quiz	quiz
6	Exam	quiz
3	Supplement	supplement
13	graded lti	Lti

En la siguiente sección se analiza como estas actividades de aprendizaje pueden ser identificadas dentro del patrón de URLs de las páginas que el estudiante visita en la plataforma.

3 Mapeo de actividades a partir de la navegación en los URLs

La mayoría de las herramientas recogen datos sobre el uso de los recursos de aprendizaje, evaluaciones, interacción social y el tiempo invertido por los estudiantes [6], datos a los cuales generalmente tienen acceso en los logfiles de las herramientas. El mapeo de las acciones de los estudiantes en la plataforma de aprendizaje permite dar seguimiento a las actividades que realizan los estudiantes durante su proceso de aprendizaje. En este estudio se analizaron los patrones de los URLs que visita el estudiante durante una sesión de estudio en la plataforma Coursera. Esta técnica de seguimiento permite recoger datos en tiempo real para ofrecer retroalimentación a los estudiantes.

En Coursera, a diferencia de otras plataformas como edX o MiriadaX, el URL visitado por el estudiante contiene información suficiente para dar seguimiento a las acciones que realizan los estudiantes y posteriormente asociar estas acciones con las actividades de aprendizaje definidas en el curso. Durante el estudio se analizó la estructura de los URLs de 13 cursos MOOCs ofrecidos por la Pontificia Universidad Católica de Chile en la plataforma Coursera. Los cursos abordan diferentes temáticas tales como: programación, educación, gestión de proyectos, física, sistemas de transportes y gestión empresarial. Esta diversidad de temáticas permite analizar diferentes opciones de navegación, según las actividades de aprendizaje definidas en el diseño de cada curso. La tabla 2 presenta las diferentes opciones de navegación que tiene el estudiante en la plataforma Coursera. Cada una de las opciones de navegación está asociada a un URL de ejemplo, tomado de nuestros cursos. Cada URL se analiza y se divide en componentes según la información que representa cada componente del URL.

El primer componente identificado en el URL es la raíz, que hace referencia al dominio de la plataforma Coursera: “*https://www.coursera.org*”. A partir de la raíz se realiza toda la navegación en Coursera, el estudiante puede realizar actividades como: registrarse en la plataforma, iniciar sesión, buscar cursos, entre otros. Sin embargo, desde esta ruta el estudiante no tiene opción de realizar actividades de aprendizaje. El segundo componente identificado es el componente “*learn*”, el cual indica que el estudiante se encuentra dentro de un curso. Este componente por sí solo no forma ninguna ruta de navegación, y requiere del tercer componente que es el “*nombre-del-curso*”, este es un nombre corto único asignado por Coursera a cada curso. La unión de los tres primeros componentes identifica el curso en el cual se encuentra el estudiante. Para facilitar el análisis de los componentes del URL que permiten dar seguimiento a las acciones de los estudiantes, los tres primeros componentes se agrupan en un solo componente, identificado como “*ruta-curso*” y el cuál será utilizado de aquí en adelante para hacer referencia a estos tres componentes. La composición de *ruta-curso* es la siguiente:

ruta-curso = raíz + learn + nombre-del-curso

Un ejemplo de este componente es:

<https://www.coursera.org/learn/electrones-en-accion>

A partir de esta ruta (*ruta-curso*), es posible extraer información para dar seguimiento a las actividades que los estudiantes realizan en cada curso. En el análisis realizado a los URLs se identificaron 10 opciones de navegación, considerando los componentes del URL que se agregan a partir del componente *ruta-curso*. Como parte del URL se identifica el componente “*código-recurso*”, el cual permite reconocer a cuál actividad de aprendizaje está accediendo el estudiante. Las actividades como video lecciones, exámenes, revisión de sus calificaciones, entre otras, siguen una estructura simple en sus componentes. Sin embargo, la actividad dirigida a la revisión por pares (*peer*) sigue una estructura más compleja. La revisión por pares se utiliza para que un grupo de estudiantes revise la tarea de otro estudiante. La revisión por pares se descompone en tres tipos de subactividades: **1)** resolver la tarea asignada; **2)** revisar la tarea de un compañero y **3)** revisar la retroalimentación dada por los compañeros. Cada subactividad crea un nuevo URL a partir del URL de la actividad de revisión base. La estructura del URL de una actividad de revisión por pares es la siguiente:

ruta-curso/tipo-recurso/código-recurso/nombre-recurso

Ejemplo:

<https://www.coursera.org/learn/aulaconstructivista/peer/ClyHs/creacion-de-un-juego-para-fomentar-lo-que-no-esta-en-el-curriculum>

En esta ruta el estudiante tiene acceso a la descripción de la tarea a realizar y la opción de enviar su tarea a revisión por parte de sus compañeros. Cuando el estudiante accede a la subactividad de revisar la tarea de otro compañero, la estructura del URL se compone a partir del URL de la actividad de revisión, se agrega el componente “*review*” y el código de la tarea que se encuentran revisando. Por ejemplo,

<https://www.coursera.org/learn/aulaconstructivista/peer/ClyHs/creacion-de-un-juego-para-fomentar-lo-que-no-esta-en-el-curriculum/review/HYKZzFq4EeeqKgpTZZjjFg>

Cuando el estudiante accede a la subactividad correspondiente a la revisión de retroalimentación dada por los compañeros, la estructura del URL se compone a partir del patrón del URL de la actividad de revisión, y se agrega el componente “*give-feedback*”, una estructura similar a la subactividad de *review*. Por ejemplo,

<https://www.coursera.org/learn/aulaconstructivista/peer/ClyHs/creacion-de-un-juego-para-fomentar-lo-que-no-esta-en-el-curriculum/give-feedback>

Los foros de discusión, también se componen de una estructura más compleja. En Coursera se definen dos tipos de foros: **1)** foros de discusión y **2)** foros semanales. Los foros de discusión abordan temas generales al curso y no están asociados a un tema o semana en particular. Los foros semanales son específicos a los temas abarcados en una semana (módulo) en particular. En la estructura del URL de los foros de discusión utiliza el componente “*forums*” y el código del recurso, ejemplo:

https://www.coursera.org/learn/electrones-en-accion/discussions/forums/r11o_2voEeWdPxJ-h_8T7w

Tabla 2. Composición de URLs en Coursera y actividades

Actividad	URL ejemplo	Composición	Descripción
1 Página de inicio	https://www.coursera.org/learn/electrones-en-accion/home/welcome	Ruta-curso/home/regla-navegación	La regla identificada es “welcome”
2 Calificaciones	https://www.coursera.org/learn/electrones-en-accion/home/assignments	Ruta-curso/home/regla-navegación	La regla identificada es “assignments”
3 Información del curso	https://www.coursera.org/learn/electrones-en-accion/home/info	Ruta-curso/home/regla-navegación	La regla identificada es “info”
4 Vista general módulo o semana	https://www.coursera.org/learn/electrones-en-accion/home/week/1	Ruta-curso/home/week/regla-navegación	La regla identificada corresponde al número de semana [1,2,3,4].
5 video	https://www.coursera.org/learn/electrones-en-accion/lecture/VX8ON/video-de-bienvenida	Ruta-curso/tipo-recurso/código-recurso/nombre-recurso	Un video es un recurso tipo “lecture”
6 Examen	https://www.coursera.org/learn/electrones-en-accion/exam/IlyXY/acerca-de-la-electricidad	Ruta-curso/tipo-recurso/código-recurso/nombre-recurso	Una evaluación formativa e examen es un recurso tipo “exam”
7 Foros	https://www.coursera.org/learn/electrones-en-accion/discussions/	Ruta-curso/tipo-recurso/	Un foro es un recurso tipo “discussions”.
8 Lecturas, pdf, links, imágenes, etc.	https://www.coursera.org/learn/experimentar/supplement/qA23f/monitorea-aprendizaje	Ruta-curso/tipo-recurso/código-recurso/nombre-recurso	Un recurso como una lectura es un recurso tipo “supplement”.
9 Revisión pares	https://www.coursera.org/learn/experimentar/peer/	Ruta-curso/tipo-recurso/	Una actividad de revisión por pares es un recurso tipo “peer”.
10 Tarea de programación	https://www.coursera.org/learn/experimentar/programming/TgURI/tarea-de-programacion-en-python	Ruta-curso/tipo-recurso/código-recurso/nombre-recurso	Una actividad de revisión por pares es un recurso tipo “programming”.

Por otro lado, el URL de los foros semanales utilizan el componente “weeks” seguido por el número de semana, ejemplo:

<https://www.coursera.org/learn/electrones-en-accion/discussions/weeks/1>

Los dos tipos de foros se componen de hilos de discusión, donde cada hilo tiene su propio código. Un ejemplo para un hilo de discusión es el siguiente:

<https://www.coursera.org/learn/electrones-en-accion/discussions/weeks/1/threads/6qmEoSvXEeabrBIWZEFM5A>

Para tener una visión más clara de la estructura y componentes de los URLs de navegación en Coursera, en la Figura 2 se muestra la jerarquía de niveles de los componentes, donde el dominio de la plataforma se utiliza como la raíz de la jerarquía.

A partir de las 10 opciones de navegación (Tabla 2) extraídas de los URLs de la plataforma Coursera se definieron 4 categorías según el rol que desempeña la actividad en el curso (Tabla 3). Los niveles de actividades por debajo del componente “home” y “supplement” se clasifican como actividades generales de aprendizaje, las cuales complementan o brindan información complementaria a las actividades principales del curso. Las video lecciones son el principal recurso utilizado en las plataformas MOOCs para entregar el fundamento teórico del curso. Este tipo de actividad ofrece información útil para analizar la interacción del estudiante con el contenido del curso. Por otro lado, los foros ofrecen información sobre el comportamiento de los estudiantes en actividades de discusión. Finalmente, todas las actividades de evaluación son agrupadas en la categoría de evaluaciones.

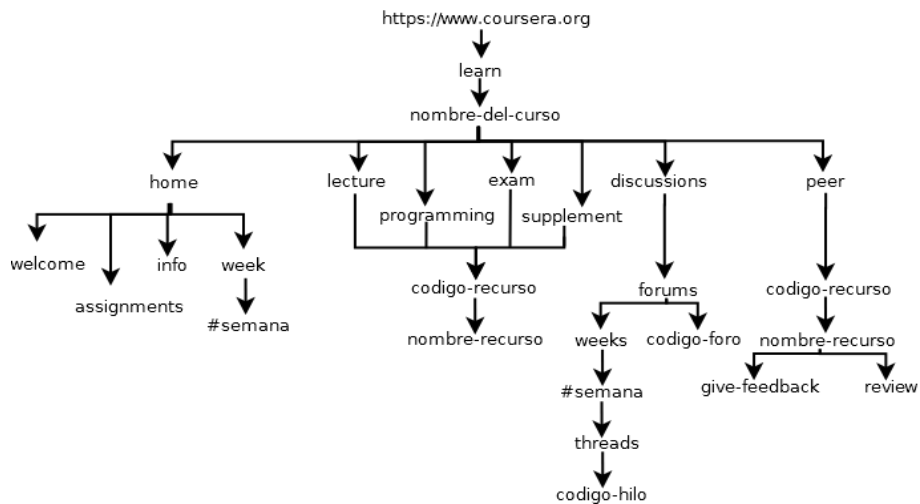


Fig. 2. Jerarquía de niveles de los componentes de los URLs de Coursera.

El seguimiento y recolección de datos basados en las 4 categorías de actividades, identificadas desde la navegación del estudiante en los URLs de la plataforma MOOCs, permite aplicar técnicas de analítica de aprendizaje y ofrecer retroalimentación en tiempo real a los estudiantes sobre su proceso de aprendizaje. Las 4 categorías de actividades se asocian a 4 de las categorías de indicadores señaladas por Schwendimann et al. [7]: **1)** las acciones que realizan los estudiantes en la plataforma; **2)** el contenido involucrado en el proceso de aprendizaje; **3)** el contexto en que tuvo lugar el aprendizaje (mediante el uso de tecnologías móviles) y **4)** la interacción social con otros estudiantes. Sin embargo, una limitación de los datos recolectados es que no es posible ofrecer retroalimentación de los resultados obtenidos en las evaluaciones. La información que se extrae de los URLs también puede ser usada por los docentes del curso para dar seguimiento al comportamiento de los estudiantes, y analizar, por ejemplo, qué tipo de material utilizan más o cuánto tiempo invierten en cada categoría de las actividades.

Sin embargo, mediante el seguimiento de los URLs no es posible determinar si el estudiante completa o no una actividad. La herramienta NoteMyProgress [8] utiliza las categorías de indicadores extraídos de los URLs de la navegación de los estudiantes para retroalimentar en tiempo real a los estudiantes con datos sobre su tiempo en las actividades de aprendizaje, tiempo fuera de la plataforma de aprendizaje, actividades realizadas, sesiones y tiempo de las sesiones realizadas, y tiempo en cada categoría de aprendizaje. Este tipo de indicadores también han sido utilizados en otras iniciativas de analítica del aprendizaje [9] [10].

Tabla 3. Categorías de actividades según URLs.

Categoría	URL en esta categoría	Descripción
Complemento	https://www.coursera.org/learn/electrones-en-accion/home/welcome	Actividades que el estudiante realiza como complemento del curso, tales como: actividades de información, recorrido por el contenido del curso, actividades fuera de la plataforma que sirven para complementar el contenido del curso, entre otros.
	https://www.coursera.org/learn/electrones-en-accion/home/assignments	
	https://www.coursera.org/learn/electrones-en-accion/home/info	
	https://www.coursera.org/learn/electrones-en-accion/home/week/...	
	https://www.coursera.org/learn/experimentar/supplement/...	
Video	https://www.coursera.org/learn/electrones-en-accion/lecture/...	Se refiere a las video lecciones
	https://www.coursera.org/learn/experimentar/p eer/...	Actividades de evaluación que realiza el estudiante.
Evaluación	https://www.coursera.org/learn/experimentar/p eer/.../give-feedback	
	https://www.coursera.org/learn/experimentar/p eer/.../review	
	https://www.coursera.org/learn/electrones-en-accion/exam/...	
Foro	https://www.coursera.org/learn/experimentar/p rogramming/...	Foros de discusión en que participa el estudiante.
	https://www.coursera.org/learn/electrones-en-accion/discussions/	

4 Conclusiones

En este estudio se analizaron los patrones que componen la estructura de los URLs que visitan los estudiantes durante una sesión de estudio en la plataforma Coursera. Mediante el análisis de los URLs se establece una relación entre la navegación del estudiante y las actividades de aprendizaje planteadas en el curso. Esta relación permite dar seguimiento a las trazas de aprendizaje y ofrecer retroalimentación en tiempo real. Los resultados de esta investigación sirven como guía para futuras investigaciones dirigidas al análisis del comportamiento de los estudiantes en esta plataforma MOOC y así aplicar técnicas de analítica del aprendizaje.

La plataforma Coursera ofrece suficiente información en la estructura de los URLs para identificar las actividades que realiza el estudiante. Los componentes de los URLs permiten identificar el tipo de actividad con que interactúa el estudiante (ver video, ingresar a una evaluación, ingresar a una video lectura), y el identificador o código de la actividad. Esta información en combinación con la información contenida en los paquetes de datos que ofrece la plataforma puede ser útil para el desarrollo de iniciativas de analítica de aprendizaje que busquen retroalimentar al estudiante en tiempo real. Con los datos recolectados se puede ofrecer retroalimentación al estudiante sobre su tiempo invertido en las diferentes actividades, interacción con las actividades, su efectividad en el uso del tiempo mientras estudia, su desempeño con respecto a otros estudiantes, entre otras alternativas.

El análisis de la información sobre las actividades de los estudiantes en la plataforma Coursera permite identificar 4 de los 6 tipos de identificadores propuestos por Schwendimann et al. [7], utilizados comúnmente en investigaciones dirigidas a la analítica del aprendizaje. Sin embargo, estos indicadores pueden ofrecer información útil para crear conciencia en los estudiantes sobre su proceso de aprendizaje.

En el estudio no se analiza el comportamiento de los estudiantes en el curso, solo se analizan los patrones obtenidos en los URLs, la utilidad de la información recolectada dependerá del objetivo de cada investigación.

Agradecimientos

Este estudio ha sido financiado por el FONDECYT (11150231), los proyectos Europeos Erasmus + MOOC-Maker (561533-EPP-1-2015-1-ESEPPKA2-CBHE-JP) y LALA (586120-EPP-1-2017-1-ES-EPPKA2-CBHE-JP) y CONICYT Doctorado Nacional 2016/21160081, CONICYT Doctorado Nacional 2017/21170467, la Universidad de Costa Rica (UCR) y la Universidad de Cuenca (Ecuador).

References

- [1] Ferguson, R.: Learning analytics: drivers, developments and challenges. *Int. J. Technol. Enhanc. Learning* 4 (5/6), 304–317 (2012).
- [2] Kulkarni, C.E., Bernstein, M.S., Klemmer, S.R.: Peerstudio: rapid peer feedback emphasizes revision and improves performance. In: *L@S 2015*, pp. 75–84 (2015).

- [3] Davis, D., Triglianios, V., Hauff, C., Houben, G.J.: SRLx: A Person-alized Learner Interface for MOOCs. In *Lifelong Technology-Enhanced Learning, EC-TEL 2018. Lecture Notes in Computer Science*, vol 11082, pp. 122-135. Springer, Cham (2018)
- [4] Shah, D.: A Product at Every Price: A Review of MOOC Stats and Trends in 2017 — Class Central. 2018. [Online]. Available: <https://www.class-central.com/report/moocs-stats-and-trends-2017/>. [Accessed: 23-Mar-2018].
- [5] Shah, D.: Massive List of MOOC Providers Around the World. Cl. Cent., <https://www.class-central.com/report/mooc-providers-list/>. [Accessed: 02-Nov-2017].
- [6] Bodily, R., Verbert, K.: Trends and issues in student-facing learning analytics reporting systems research. In: *Proceedings of Seventh International Learning Analytics & Knowledge Conference - LAK 2017*, pp. 309–18 (2017).
- [7] Schwendimann, B.A., et al.: Perceiving learning at a glance: a systematic literature review of learning dashboard research. *IEEE Trans. Learn. Technol.* 10, 30–41 (2017).
- [8] Pérez-Álvarez, R., Pérez-Sanagustín, M. and Maldonado-Mahauad, J. J.: NoteMyProgress: Supporting learners' self-regulated strategies in MOOCs. *EC-TEL 2017. LNCS*, vol. 10474, pp. 517–520. Springer, Cham (2017).
- [9] Guerra, J., Hosseini, R., Somyurek, S. and Brusilovsky, P.: An Intelligent Interface for Learning Content: Combining an Open Learner Model and Social Comparison to Support Self-Regulated Learning and Engagement. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*, pp. 152–163. (2016)
- [10] Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Leony, D., and Delgado Kloos, C.: ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. *Comput. Human Behav.* 47, 139–148 (2015).