

Designing dashboards for students and instructors in a Sakai supported face-to-face learning environment

Haro-Valle, Valeria¹, Benlloch-Dualde, José-V. ¹, Lemus-Zúñiga, Lenin y Maldonado-Mahauad, Jorge J.²

¹ Universitat Politècnica de València, Valencia 46022, Spain
vahaval@inf.upv.es, jbenlloc@disca.upv.es, lemus@upv.es

² Universidad de Cuenca, Ecuador
jjmaldonado@uc.cl

Abstract. A very recent review of the literature in the field of Visual Learning Analytics states that the use of these techniques is quite frequent in blended learning or online learning environments, including MOOCs. However, this is not the case in classroom learning environments. In this context, the paper aims at studying how using Visual Learning Analytics can contribute to a better understanding of the educational processes in face-to-face educational contexts, supported by a Sakai-based Virtual Learning Environment. Considering that the institutional platform reports are only available for teachers, the main objective is the design and implementation of a learning dashboard that could help students in their learning process. To accomplish that, it should integrate data from different sources, generating easy to understand visual representations. Additionally, the same data sources will be used to develop a dashboard for instructors that could help them to provide formative feedback to their students, or to improve the teaching materials they use. In order to create the dashboards, visualization tools such as *Tableau* and *QlikSense* have been initially employed. However, it was decided to use the java script library *D3.js*, as it allows us to create any imaginable visualization and because of the interactivity it offers. To conclude, some preliminary results are discussed, and further research is outlined.

Keywords: Visual Learning Analytics, Learning Dashboard, Face-to-Face Learning Environment, Sakai.

1 Introducción

En la actualidad la información es, sin duda, uno de los pilares fundamentales en nuestra sociedad y, cada vez más, las técnicas de análisis de datos se utilizan como ayuda a la toma de decisiones en todo tipo de organizaciones y contextos. Sin embargo, a pesar de que en el ámbito académico los volúmenes de información han crecido enormemente por la incorporación de las TIC y, sobre todo, por los entornos virtuales de aprendizaje, el uso de dichas técnicas tiene todavía un gran potencial.

La analítica de datos en el ámbito educativo tiene una vida relativamente corta, pues no se empieza a utilizar el término *Learning Analytics (LA)* hasta el año 2009. Pocos

años después se acuña una definición bastante aceptada del mismo, como: “la medición, recopilación, análisis y realización de informes de datos sobre los alumnos y sus contextos, con el propósito de entender y optimizar el aprendizaje y los entornos en los que se produce” [1].

No menos importantes que las técnicas de análisis son las estrategias para visualizar esos conjuntos de datos y cómo el tipo de representaciones escogidas contribuyen a facilitar su interpretación. Esta disciplina se conoce como *Visual Analytics* (VA) y ha sido definida como “la ciencia del razonamiento analítico facilitada por las interfaces visuales interactivas” [2]. Si los profesores tuvieran acceso a visualizaciones efectivas de sus datos educativos, podrían usarlos para proporcionar retroalimentación formativa a sus estudiantes o para mejorar los materiales didácticos que emplean. Del mismo modo, si los estudiantes pudieran tener acceso a este tipo de visualizaciones, se podría favorecer el desarrollo de sus habilidades de autorregulación del aprendizaje y ayudarles a conseguir sus objetivos.

En un trabajo muy reciente [3], se realiza una revisión sistemática de la literatura en la intersección de las áreas citadas anteriormente y se introduce el término *Visual Learning Analytics* (VSA) como “el uso de herramientas y métodos computacionales para la comprensión de fenómenos educativos a través de técnicas de visualización interactiva”. De los 52 trabajos revisados en el estudio, tan solo tres correspondían a contextos de enseñanza presencial, siendo los más frecuentes los entornos híbridos (*blended learning*), los enteramente on-line o los MOOC. En esta línea, este trabajo se propone estudiar cómo el uso de técnicas de VLA puede contribuir a comprender mejor los procesos educativos en cursos de enseñanza presencial apoyados en el uso de la plataforma institucional Poliformat [4], basada en Sakai. En este planteamiento, resulta de gran interés un estudio anterior [5] que aborda la relación entre los datos de interacción en la plataforma con el rendimiento académico, ya que analiza también este mismo tipo de contextos educativos, también llamados *VLE-supported F2F courses*.

Tal y como apunta un reciente informe del Departamento de Educación de los Estados Unidos [6], la información proporcionada a estudiantes y profesores sobre el nivel de logro de los objetivos de aprendizaje es más valiosa si está disponible en un lugar de fácil acceso. Con este fin, aboga por introducir los *dashboards* de aprendizaje, que pueden utilizarse para integrar información de distintas fuentes generando representaciones visuales fáciles de entender.

El principal objetivo de este trabajo es el diseño e implementación de un *dashboard* en un entorno de enseñanza presencial donde el estudiante pueda, de manera sencilla, interpretar las gráficas que se le presentan para monitorizar su aprendizaje durante el curso académico. Por otro lado, se creará también un *dashboard* para el profesor de modo que pueda, a partir de los datos presentados, identificar a los alumnos en riesgo, observar el acceso a los materiales o la entrega de las tareas propuestas, entre otras, y así poder tomar alguna medida correctora.

El resto del trabajo se estructura como sigue: la sección 2 describe la arquitectura del sistema y los bloques que la forman; la sección 3 expone los criterios seguidos en el diseño del *dashboard*; la sección 4 introduce los resultados preliminares y, por último, la sección 5 presenta las conclusiones.

2 Arquitectura

Para la implementación del *dashboard* se considera la arquitectura descrita en la Fig. 1, donde se definen 3 bloques. En el primero, se identifican las fuentes de datos, en el segundo se realizan procesos de extracción, limpieza y carga para preparar los datos que serán la entrada del bloque 3 y, en este último, se diseña e implementa el *dashboard*.

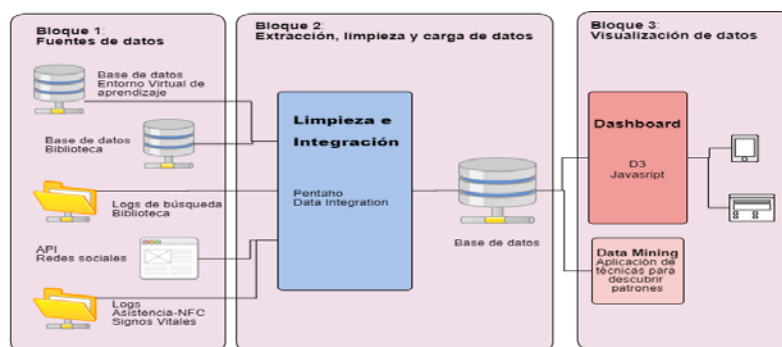


Fig. 1. Arquitectura del *dashboard*

2.1 Fuentes de datos

Dado que el contexto de nuestro estudio es una enseñanza presencial soportada por el uso de un entorno virtual de aprendizaje, los datos proporcionados por la plataforma van a ser la principal fuente de datos. En nuestro caso, la plataforma institucional, y más concretamente su herramienta “Estadísticas”, permite generar informes sobre las interacciones de los usuarios con aquella, para cada uno de los cursos donde de utiliza. Para cada informe se puede definir el tipo de actividad (visitas, eventos, acceso a recursos, tiempo de estancia), el intervalo temporal al que se refiere, los usuarios a monitorizar y cómo se mostrarán los resultados. Una vez generados los datos, se exportan en ficheros CSV o XLS. Si bien la cantidad y calidad de los datos que se pueden obtener en estos informes son considerables, es importante señalar que siempre están restringidos a las opciones implementadas en la herramienta y que, por tanto, no se tiene un acceso completo a la base de datos que almacena los *logs* de los usuarios que acceden a la plataforma.

Por otro lado, se utilizan también los ficheros de calificaciones de las asignaturas, igualmente en formato CSV, que almacenan las notas de los diferentes actos de evaluación y que serán de gran utilidad para relacionar la información obtenida de la plataforma con la de rendimiento académico.

Las fuentes de datos anteriores se podrían complementar con muchas otras, algunas genéricas y otras dependientes del enfoque utilizado en cada asignatura, tal y como se apunta en diversos estudios [7, 8]. Entre las primeras cabe considerar, por ejemplo, los ficheros de asistencias a las sesiones presenciales, los accesos a las bases de datos de la biblioteca o incluso las bases de datos donde se guardan los expedientes de los

estudiantes y que incluyen, entre otros, datos demográficos, notas de acceso, opciones de matrícula, rendimiento en cursos anteriores. Aunque el acceso a estos datos puede ser, en no pocas ocasiones, bastante difícil de obtener, las posibilidades que abren son enormes. Respecto a las fuentes de datos específicas, en algunos cursos se utilizan aplicaciones tales como *Socrative*¹ o *Kahoot*², por citar dos muy populares, para generar cuestionarios. Las respuestas de los estudiantes se recogen en una serie de informes que podrían ser igualmente de gran utilidad. En otros cursos se hace uso de redes sociales como Twitter, Facebook, Instagram u otras. Las respectivas API de estas redes podrían proporcionar igualmente fuentes de datos a incorporar a nuestro sistema. Dependiendo de las herramientas utilizadas en cada curso, se pueden integrar distintas fuentes de datos.

2.2 Extracción, transformación y carga (ETL)

Los procesos ETL (*Extract, Transform and Load*), permiten copiar los datos de una fuente y cargarlos en otra. En la primera fase, denominada extracción, se toman los datos de diferentes fuentes, ya sean bases de datos, archivos de texto, hojas de cálculo, entre otras. Posteriormente, en la fase de transformación, se realiza una limpieza de datos y se modifican para estandarizar los datos en un determinado formato. Finalmente, estos datos serán cargados en otro repositorio durante la fase de carga.

Utilizando la herramienta *Pentaho Data Integration* o también conocida como *Kettle*³, se realizan varios procesos ETL con la finalidad de preparar los datos que serán mostrados en el *dashboard*. En la Fig. 2 se expone el proceso ETL de los datos de la asignatura Tecnología de Computadores (TCO), de primer curso del Grado de Ingeniería Informática de la Universitat Politècnica de València, donde se recogen los informes de los eventos realizados por cada alumno, el tiempo por día que el estudiante ha accedido al sistema, los recursos a los que accede y las notas finales obtenidas por los estudiantes. Estos datos se recogen en archivos Excel y se integran mediante el identificador del estudiante, además, se realizan transformaciones de datos como el reemplazo de los datos nulos por “0” y la extracción del identificador del grupo al que pertenece el estudiante de una cadena de caracteres. Y durante el paso final se almacena los datos en un nuevo archivo.

Las variables disponibles tanto para la asignaturas de Grado TCO y Fundamentos de Computadores (FCO), como para Información Multimedia en entornos Multidispositivo (IMM), esta última del master de Gestión de la Información (MUGI) son: el identificador del estudiante, asignatura, número de eventos realizados en la plataforma, fecha del evento realizado por el estudiante, tiempo de conexión a la plataforma, número de clases a las que asiste, recursos a los que accede y notas adquiridas en los diferentes actos de evaluación. A partir de estos datos se crearán indicadores para el *dashboard*.

¹<https://www.socrative.com/materials/SocrativeUserGuide.pdf>

²<https://kahoot.com/blog/2017/02/20/download-evaluate-kahoot-results-data/>

³<https://www.hitachivantara.com/en-us/products/big-data-integration-analytics/pentaho-data-integration.html>

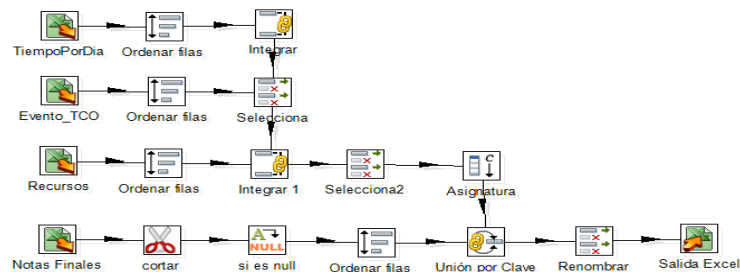


Fig. 2. Proceso ETL para preparar los datos de la asignatura TCO

2.3 Visualización de datos

Stephen Few define un *dashboard* como una visualización de la información más importante para lograr uno o más objetivos; información consolidada y organizada en una sola pantalla por lo que puede ser monitoreada de un vistazo [9].

Se han desarrollado *dashboards* para los diferentes niveles y tipos de educación, con diferentes tecnologías y diferentes enfoques. Con la popularidad que ha tenido la educación online en los últimos años, se han desarrollado las plataformas MOOC (*Massive open online courses*) para dar soporte durante el proceso de aprendizaje, así como también herramientas como *NoteMyProgress*, que permite a los usuarios del navegador Google Chrome integrarse con las plataformas MOOC para recolectar datos y facilitarle un *dashboard* al estudiante, que permite analizar información de la actividad en su aprendizaje [10].

Otra herramienta conocida como *StepUp* está dirigida principalmente a los profesores en un contexto de aprendizaje colaborativo asistido por computadora, y también facilita una versión móvil dirigida a los estudiantes para brindar información de los esfuerzos que ha realizado durante la semana y el esfuerzo general hasta la fecha [11].

Como estas herramientas mencionadas, se han desarrollado otras, como *LOCO-Analyst* [12], *GLASS* [13], *SAM* y *CourseSignal* [14] que analizan los accesos, el uso de los contenidos, mensajes y resultados del rendimiento del estudiante. Además, otras herramientas como *Student Success System* [15], *SNAPP* y *Narcissus* extraen datos de redes sociales para integrarlos con datos académicos y lograr de esta manera un análisis integral del estudiante [11].

Las herramientas anteriores se han desarrollado tanto en versiones *Web* y móvil, lo que permite que la herramienta esté al alcance de los estudiantes y profesores a través de los dispositivos que utilizan diariamente. Para el desarrollo del presente proyecto se considera la implementación de una versión *Web* multidispositivo que mediante la utilización de mensajes en formato JSON se lleve a cabo la comunicación entre el lado servidor (fuente de datos) y el lado cliente (herramienta de visualización).

3 Diseño del *dashboard*

Los *dashboards* utilizan diferentes técnicas de visualización ya sea gráficos de barras, gráficos pastel, tablas, nube de palabras, líneas de tiempo, gráficos de dispersión, entre otros, que son seleccionados apropiadamente dependiendo de la información que van a representar, con el objetivo de facilitar la abstracción de la información tanto al profesor como al estudiante.

Para la creación del *dashboard*, Stephen Few menciona 13 errores comunes en el diseño que debemos evitar, los mismos que se presentan a continuación [9]:

1. Exceder los límites de una sola pantalla
2. Proveer el contexto inadecuada de los datos
3. Visualización con excesivo detalle o precisión
4. Elegir una medida deficiente
5. La elección de gráficos inapropiados
6. Incluir alguna complejidad sin sentido
7. Uso de medios de visualización mal diseñados
8. Mostrar datos cuantitativos inexactos
9. La organización pobre de los datos
10. Destacar los datos importantes erróneamente o no resaltarlos en absoluto
11. Saturación o con demasiada decoración
12. El mal uso o el uso excesivo del color
13. Diseñar una visualización poco atractiva

Para diseñar el *dashboard* se tomaron en cuenta estas recomendaciones y se analizaron diversas opciones de herramientas de visualización de datos empleadas en el ámbito de la inteligencia de negocios y *big data*. Con el fin de ensayar distintas aproximaciones hasta llegar a definir el modelo de *dashboard* propuesto, se utilizaron inicialmente algunas herramientas disponibles en el mercado, evitando así el tiempo necesario para los correspondientes desarrollos. En particular, se optó por las aplicaciones *Tableau*⁴ y *QlikSense*⁵.

Tableau es un software comercial que facilita la visualización de datos de forma interactiva, lo cual le ha hecho popular en el campo empresarial. De forma similar, *Qlik-Sense* facilita la creación de gráficas para realizar el análisis de los datos y la creación de *dashboards* interactivos.

Por otro lado, un estudio comparativo entre *Tableau* y *D3.js*⁶, explica que se puede utilizar *Tableau* en el caso de que no se tengan conocimientos de programación, ya que brinda al usuario una interfaz gráfica muy intuitiva para realizar análisis y exploración de datos. También se afirma que el uso de *D3.js* es más apropiado cuando hay requisitos de gráficos complejos o altos requisitos de interactividad [16].

Con el fin de tener mayor flexibilidad en la creación del *dashboard* junto con mayores opciones de interactividad, para el modelo final se ha optado por un desarrollo

⁴ <https://www.tableau.com/>

⁵ <https://www.qlik.com/es-es/products/qlik-sense>

⁶ <https://d3js.org/>

propio, utilizando la librería java script *D3.js*. Esto permitirá que los estudiantes puedan acceder fácilmente a un *dashboard* personalizado y escalable.

4 Resultados

Después de ensayar los distintos tipos de gráficos ofrecidos por las herramientas citadas en la sección anterior, así como sus respectivas combinaciones, se ha optado por el modelo de *dashboard* que se presenta en la Fig. 3, que se describe a continuación. En la sección superior izquierda, una gráfica de tipo velocímetro que indica al estudiante mediante colores, el estado que presenta en una determinada asignatura, donde el rojo significa riesgo, el amarillo precaución y el verde significa que el estudiante tiene un buen rendimiento académico. Además, se muestra la nota que ha obtenido el estudiante en las dos últimas tareas o exámenes. También se incluye un gráfico que muestra la popularidad de los recursos que están al alcance del estudiante, y otro gráfico de pastel para que el estudiante conozca el número de eventos, según el tipo de evento a los que accede. En la sección de la derecha se tienen dos gráficas con el tiempo de acceso a la plataforma para examinar el ritmo que lleva el estudiante mensualmente y durante las últimas semanas. Incluir adicionalmente un indicador de la media del tiempo de actividad de su clase es clave para que el estudiante pueda medir si está comprometido en el aprendizaje de la asignatura como sus otros compañeros o, por el contrario, necesitaría una mayor implicación.

Es importante resaltar que, en la parte superior de la Fig 3, se pueden distinguir los filtros que puede realizar el usuario del *dashboard*, permitiendo la interacción del usuario con la herramienta y la generación de consultas por asignatura, año o mes, y donde, únicamente el perfil profesor, podrá realizar el filtro por un estudiante determinado o un grupo de estudiantes.

Además, para el perfil profesor se incluyen las gráficas de barra mostradas en la Fig 4. donde, basándonos en las notas que ha obtenido el estudiante previamente, el profesor se mantiene informado durante todo el periodo académico de qué estudiantes tienen mayor riesgo en la asignatura, al tiempo que le permite identificar aquellos con mejores resultados. Por otra parte, le permite analizar la correlación entre asistencia y calificaciones y monitorizar la actividad en la plataforma.

El trabajo aún está en progreso, pero en análisis preliminares y exploratorios que se realizaron con los datos facilitados, comienzan a aparecer resultados interesantes. Como puede observarse en la Fig. 5, en un análisis por grupos para una asignatura de primer curso de Grado en Ingeniería Informática (TCO) existe una correlación positiva entre el tiempo de estancia del estudiante en la plataforma y la nota final obtenida, lo que apoya con datos la utilidad de la plataforma como soporte al aprendizaje.



Fig. 3. Dashboard diseñado para dar soporte al estudiante en su proceso de aprendizaje

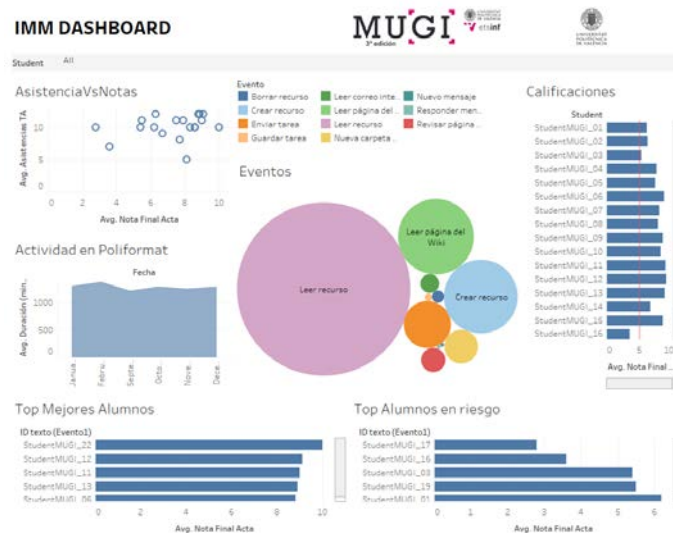


Fig. 4. Dashboard diseñado para el profesor en la asignatura IMM

Por otro lado, se han encontrado ciertas particularidades en la utilización de la plataforma Poliformat, dependiendo de la forma en que los estudiantes trabajan en cada curso. Por ejemplo, en una de las asignaturas de máster analizadas (Información multimedia en entornos multidispositivo), el profesor ofrece la posibilidad de realizar todos los trabajos y tareas del curso, tanto en equipo como individualmente. Este hecho ha causado que uno de los integrantes del grupo, apenas registre actividad en Poliformat, lo que complica los análisis y supone mostrar datos erróneos en el *dashboard*, pues las interacciones de este/a estudiante debieran asociarse a las de su compañero/a de equipo.

Para solucionar esta particularidad, cada estudiante pertenecerá a un grupo dentro de su clase, así después las estadísticas podrán ser generadas en base al identificador del grupo en lugar de usar el identificador individual del estudiante.

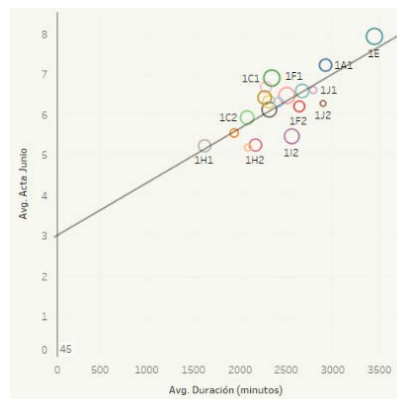


Fig. 5. Correlación entre el tiempo de acceso y la calificación

5 Conclusiones

Este trabajo ha indagado en el estudio de cómo el uso de técnicas de VLA puede contribuir a comprender mejor los procesos educativos en cursos de enseñanza presencial apoyados en el uso de entornos virtuales de aprendizaje, tales como Sakai. El objetivo principal ha sido el diseño e implementación de un *dashboard* para soporte académico que visualice, en una única pantalla, diferentes fuentes de datos mediante la combinación de distintos tipos de gráficos. El *dashboard* muestra, por ejemplo, los recursos más populares entre los estudiantes, lo que permite una mejora continua de los materiales didácticos. Por otro lado, al informar de la actividad del estudiante en la plataforma, y compararla con la del resto de compañeros, puede contribuir a disminuir la tasa de abandono porque facilita la identificación anticipada de los alumnos en riesgo y da a los estudiantes el control sobre su aprendizaje, promoviendo una cultura de autorregulación y toma de decisiones basadas en datos.

Respecto al diseño, la estrategia inicial de utilizar herramientas de visualización de datos ya disponibles en el mercado, tales como *Tableau* y *QlikSense*, mientras se definía el modelo de *dashboard*, ha resultado de gran utilidad, al evitar desarrollos innecesarios y facilitar el conocimiento de los datos mediante la exploración y el análisis. El siguiente paso es su desarrollo Web. Además, se pretenden incluir nuevas fuentes para ayudar a identificar otras particularidades. Por último, está previsto realizar una fase de evaluación para identificar la utilidad de la propuesta y recibir las recomendaciones y sugerencias de la audiencia a quien se dirige, esto es, estudiantes y profesores.

Agradecimientos

Este trabajo ha sido financiado por el máster universitario de Gestión de la Información (MUGI) de la ETSINF, en la Universitat Politècnica de València.

Referencias

1. Siemens, G., Gasevic, D.: Guest editorial-Learning and knowledge analytics. *Educational Technology & Society* 15(3), 1–2 (2012).
2. Thomas, J. J., Cook, K. A.: *Illuminating the path: The research and development agenda for visual analytics*. IEEE Computer Society (2005).
3. Vieira, C., Parsons, P., Byrd, V.: Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers & Education* 122, 119-135 (2018).
4. Mengod, R.: Poliformat, the Sakai-based on-line campus for UPV - history of a success. In: *Proceedings of the 5th Sakai Conference*, Vancouver, BC, Canada (2006).
5. Agudo-Peregrina, Á. F., Iglesias-Pradas, S., Conde-González, M.A., Hernández-García, A.: Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. *Computers in Human Behavior* 31, 542-550 (2014).
6. U.S. Department of Education: *Future ready Learning: Reimagining the role of technology in education - 2016 national education technology plan* (2016). Retrieved from <http://tech.ed.gov/files/2015/12/NETP16.pdf>.
7. Kitto, K., Cross, S., Waters, Z. and Lupton, M.: Learning analytics beyond the LMS: the Connected Learning Analytics toolkit. In: *Proceedings of the 5th International Learning Analytics and Knowledge (LAK'15) Conference*, pp. 11-15. ACM, Poughkeepsie, New York, USA (2015).
8. Buendía-García, F., Benlloch-Dualde, J.V.: Las Fuentes del Learning Analytics. Más allá de las Plataformas de Aprendizaje. In García-Peñalvo, F.J., Mendes, A.J. (eds.). *Simposio Internacional de Informática Educativa (18º. 2016. Salamanca)*, pp. 255-260. Ediciones Universidad de Salamanca, Salamanca (2016).
9. Few, S.: *Information dashboard design: The Effective Visual Communication of Data*. 1st edn. O'Reilly Media (2006).
10. Pérez-Álvarez, R., Pérez-Sanagustín, M. and Maldonado-Mahauad, J.J.: NoteMyProgress: Supporting learners' self-regulated strategies in MOOCs. In: *European Conference on Technology Enhanced Learning*, pp. 517-520. Springer, Cham (2017).
11. Park, Y. and Jo, I.: Development of the Learning Analytics Dashboard to Support Students' Learning Performance. *Journal of Universal Computing Science* 21(1), 110–133 (2015).
12. Jovanović, J. et al.: LOCO-analyst: A tool for raising teachers' awareness in online learning environments. In: *European Conference on Technology Enhanced Learning*. Springer, Berlin, Heidelberg, (2007).
13. Leony, D., Pardo, A., de la Fuente Valentín, L., de Castro, D. S., & Kloos, C. D.: GLASS: a learning analytics visualization tool. In: *Proceedings of the 2nd international conference on learning analytics and knowledge*. ACM, pp. 162-163 (2012).
14. Arnold, K. E. and Pistilli, M.D.: Course signals at Purdue: Using learning analytics to increase student success. In: *Proceedings of the 2nd international conference on learning analytics and knowledge*. ACM, pp. 267-270 (2012).
15. Essa, A. and Ayad, H.: Student success system: risk analytics and data visualization using ensembles of predictive models. In: *Proceedings of the 2nd international conference on learning analytics and knowledge*. ACM, pp. 158-161 (2012).
16. Nair L., Shetty S., Shetty S.: Interactive visual analytics on Big Data: Tableau vs D3.js. *Journal of e-Learning and Knowledge Society* 12(4), 139-150 (2016).