● Luis De La Fuente y Daniel Burgos
Logroño (La Rioja, España)

Am I doing well? A4Learning as a self-awareness tool to integrate in Learning Management Systems

¿Lo estoy haciendo bien? A4Learning como una herramienta de auto-conciencia de integración en los Sistemas de Gestión de Aprendizaje

ABSTRACT
Most current online education scenarios use a Learning Management System (LMS) as the basecamp for the course activities. The LMS offers some centralized services and also integrates functionality from third party services (cloud services). This integration enriches the platform and increases the educational opportunities of the scenario. In such a distance scenario, with the students working in different physical spatial locations, they find difficult to determine if their activity level matches the expectation of the course. A4Learning performs a daily-updated analysis of learners' activities by establishing the similarity between two given students. That is, finds students that are doing similar things in the Learning Management System. Then, the system finds and represents how similar students have similar achievements in the course. A4Learning can be integrated within the LMS to provide the students with a visual representation their similarity with others as an awareness mechanism, so that the students can determine the achievements of similar students in previous courses and estimate their own performance.

KEYWORDS
Self-awareness, Learning analytics, Learning Management System, Similarity measurement, eLearning, LMS.

SOBRE EL AUTOR/ES
Dr. Luis de la Fuente Valentín. Universidad Internacional de La Rioja (UNIR) (luis.delafuente@unir.net).
Dr. Daniel Burgos. Universidad Internacional de La Rioja (UNIR) (daniel.burgos@unir.net).

1. Introduction

Awareness provision is one of the difficulties of distance learning scenarios. In face-to-face settings, students can easily exchange their thoughts on the course and share information that might be relevant for a better performance. This way, students are aware of how much effort their peers are spending on the course and can roughly estimate to what extent they are spending more or less effort than expected. Such community awareness is hardly possible in distance scenarios, with the students working in different physical spatial locations. However, distance scenarios offer tools (such as activity tracking methods) that allow for the analysis of the students' activity and enable the construction of methods to provide the required community awareness.

In this sense, analytics techniques can be used to inspect the whole of the facts, current and historical, to make predictions about future or events, usually relating the predictions with the discovery of behaviour patterns. On this way, analytics techniques can be used for the early identification of students at risk and grade prediction, which is rarely targeted at students and instead is teacher-oriented [1] at schools and universities. For example, grade estimation is usually addressed by the 'academic analytics' research field [2] with examples in the literature such as [3], that discusses the relationship between LMSs' usage patterns and students' motivation; and [4], that analyses the correlation involvement on a course and obtained grade.

A different approach is, instead of making predictions, the use of analytics to provide awareness systems to the students, so they can access to information that enables self-assessment of learning efforts or helps them taking decisions for their learning. For example, the work presented at [5] supports resource-abundance for self-regulated learners. Another example presents visualization methods to analyse trending data in the learning context [6]. As presented at [7], the provision of awareness causes an impact on student's habits. In those cases, the information obtained from a course is pre-processed and visually presented to the students in order to allow them to self-assess their knowledge and/or performance.

The rising interest on Learning Analytics has provided new tools and methods for the awareness provision. Thanks to the tracking capability of Learning Management Systems (LMS), distance online learning is a well-suited area of application for Learning Analytics. For instance, use cases of how user activity tracking challenges are met with data mining techniques in the context of the LMS are provided at [8] and [9]. In the Learning Analytics field, leveraging human judgment is key, while automated discovery is a tool to accomplish this goal [10]. Visualization techniques are therefore a common approach to show the information to the end user, which is typically referred to as "visual analytics" [11]. The idea behind visual analytics is to let a computer program filter and pre-process the data, visually arrange it and then let the user perform an interpretation. For instance, the work presented by [12] and [13] support teachers in keeping track of students' participation with graphical representations of their activity, while [14] use visual analytics to obtain visual recommendations.

This paper presents A4Learning (Alumni Alike Activity Awareness), an awareness tool that serves as score estimation method for students and is based on a visual analytics approach and on the similarity between behavioural patterns. Students are presented with a visual representation of the achievements of students from previous courses, and a measurement of how similar they are. The students can therefore view what similar students did, observe their achievements and therefore self-assess their own progress. A4Learning uses the concept of the "nudge" proposed at [15] with the goal of increasing students' self-awareness and helping them to self-asses their progress. The web tool nature of A4Learning allows its integration into different Learning Management Systems (LMS) or other learning systems, thus empowering the cloud as learning environment.

The rest of this paper is organized as follows: next section explains the background ideas of A4Learning and explains the system with a case of use; then, Section III explains the different visual approaches designed for the tool. After that, the architecture of the tool and its relation with the cloud is depicted and, finally, the conclusions are presented.

2. System description

In order to guide students at the beginning of the course, some teachers present them with statistics from previous courses. The presented information usually relates students' achievements with their attitude towards the course. For example, a teacher might say "the ratio of successful students is higher between those students who attended the lectures than those who did not attend". This way, the students are encouraged to adopt the attitude of successful students. A4Learning is backgrounded in the idea of similarity between students, and explores the relationship between behaviour and achievement, with the goal of offering feedback to the students. Thus, A4Learning offers the students the opportunity to find the learning results of peers that behaved like them. This is achieved by tracking the students' daily activity, mainly taking place at the LMS. The collected information, stored in the shape of activity logs, is used to calculate how similar are the activity logs from different students. So, if two students produce similar activity logs, we say they are students who behave similar.

Having that similarity is calculated in relation to students from previous courses (whose learning results are already known) instead of current peers, it is possible to establish a relationship among students who behaved similar and their learning results. This relationship is then visually presented to the students, and they will be able to see and evaluate this relationship by themselves and therefore foresee their own learning results and check if they match their initial expectations in the course.

As the main goal of A4Learning is to encourage students' reflective attitude by increasing their course awareness, the information offered by A4Learning is intentionally inconclusive. This way, the students are able (and responsible) to contextualize the received information into their personal circumstances and then evaluate whether or not they are reaching their expectations in the course.

2.1. Example of use

A personalized visualization is generated for every student. To explain how A4Learning offers the awareness information, we present the following example: a student, Alice, is enrolled in a programming course, which lasts three months and that it is now on the third week. During the previous course edition (with similar duration, and the same learning activities and available tools), a monitoring system captured the students' activity and now Alice can compare her own performance with the historic records of other students. In this scenario, A4Learning works as follows:

During the first three weeks of the course, Alice daily uses the LMS tools like forum, submission systems, videos, readings, interactive learning resources, etc., and her activity is being monitored. A4Learning then compares Alice's activity pattern with students from the previous course. The result is a similarity coefficient that, for each historic student, says how similar Alice and that student are. With this, all the historic records of students are represented by small circles in a scatter plot, where the horizontal and vertical coordinates represent their final grade and the colour of the circle is selected according to the student's similarity with Alice: the more similar to Alice, the darker the colour.

The resulting visualization (see Figure 1) allows Alice to observe the final grade obtained by behaviour-alike historic students. As it can be seen in the figure, students who show similar behaviour to Alice's are clustered and located in a certain place of the scatter plot, so that Alice can estimate which her own location in that graphic would be and, therefore, she is able to foresee her own performance in the course. As Alice's activity is evaluated daily, her estimation may evolve as the course follows up.

© ISSN: 2255-1514

Figure 1: Scatter plot representation of students similarity.

Therefore, A4Learning is not estimating Alice's grade, but it is showing the grade of similar historic students. Alice knows her personal context and circumstances, so she is able to put the visualization in the proper context and reflect if she is covering her learning expectations in the course.

3. A4Learning visual representation

As the analytics done by A4Learning are visually presented to the students so they are expected to make their own interpretation of the data, the used visualization method has a major relevance for the success of the tool in real courses. Therefore, the design of the visual approach follows an iterative procedure that allows for the refinement of the idea. In particular, we elaborated a first prototype of the visual representation and presented it to multidisciplinary stakeholders. Then, we individually interviewed them and used their feedback for the refinement of the design.

3.1. First visualization

The first visual approach attempts to maximize the amount of information posed to the student. To this extent, A4Learning uses scatter plots where each point represents a different student. Such a two-dimensional approach allows for the easy visual recognition of clusters, without overloading the user with exhaustive information. The scheme assumes that students participate in two types of assessed activities: laboratory (practical) sessions, and mid-term and final examinations. Such an assumption allows the identification of the horizontal and vertical axes with these two activity-types, making it easy to calculate a given student's score. Therefore, a student's position in the graphic is given as follows:

- The horizontal axis represents the student's score during practical activities (e.g., lab or exercises).
- The vertical axis represents the student's score for theoretical activities (e.g., mid-term or final exams).

Similarity is represented with different colour intensities. Following the previous example, Alice will receive a graphic where students with similar activity are represented in dark green. In order to help with the interpretation of the scatter plot a threshold line is drawn, being the pass-fail threshold. That is, students who passed the course (i.e. who scored over 5) are at the right side of the line.

© ISSN: 2255-1514

3.1.1. Evaluation with experts and new proposal

With the goal of collecting feedback and improving the A4Learning look and feel, we evaluated the tool with a team of experts covering different knowledge fields, namely:

- Two fresh engineers, who currently work as software developers and who finished their degree not long ago.
- Two experienced software engineers who currently work in management.
- One researcher with expertise in sociology and its link to teaching and learning
- Two researchers in the field of Technology Enhanced Learning

The most representative finding was that the visualization is hard to understand. That is, too much information is presented in the scatter plot and a training session is required to understand the visualization. The expected result is that the students will ignore the visualization. A new visual approach should be easier to understand at the cost of reducing the amount of presented information, so the student is not overloaded with the represented information. As a reaction to this analysis, we propose the following second version of the tool based on different visual representations, with different levels of complexity.

A prototype has been therefore designed as follows: the student will receive a simpler version of the visualization, but she will still be able to access the scatter plot and explore detailed information.

The first visualization presented student will be based on the average measurement of similarities. That is, the historic students are classified according to their final score in the course (e.g. a group of students with less than 3 points, another group with students with more than 3 but less than 5 points, etc.) and then the system calculates the average similarity in each of these groups. The tool then draws these groups as areas in a 0-to-10 score line and select the colour of each area according to the average similarity (again, the more similarity, the darker the colour). This simplified visual approach prevents the students from being overloaded with information, at the same time that keeps offering awareness to the students. The resulting prototype is shown at Figure 2.

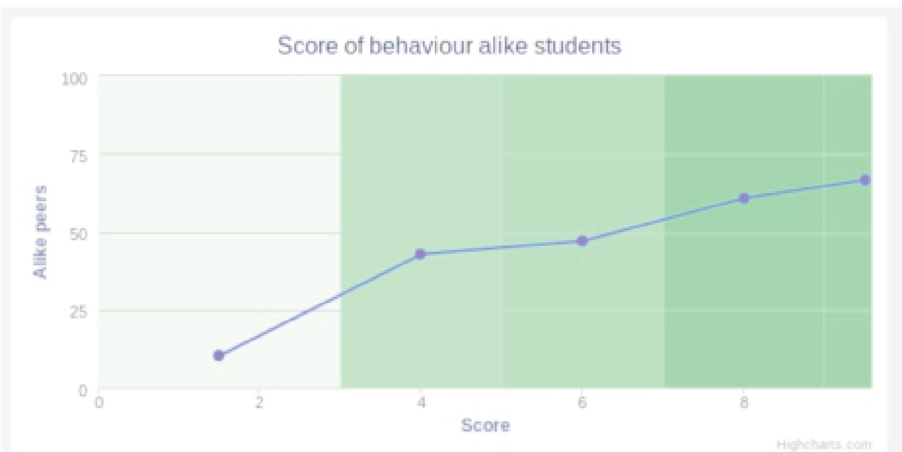

Figure 2: Average-based representation of students' similarity.

With this simpler visual approach, the student is still able to self-assess her performance in the course. However, as the available information has been reduced, the student might not be able to contextualize the received information in her personal context. Therefore, the student might make a wrong interpretation of the graphic and think that she is doing better (or worse) than she is actually doing. To avoid this, the scatterplot is still available, so the student can explore historic students in more detail.

Additionally, a new version of the scatter plot will be produced with enhanced information about similarity.

© ISSN: 2255-1514

In particular, the students will be able to find an answer to the following question: "what makes this student similar to me?". Therefore, they will be able to interactively find details (on demand) of the historic students' activity, so they will better understand what similarity means in their particular situation.

Finally, the integration with the LMS will be reduced to a lighter version of the visualization, being just an item of information that is offered to the student on the home page of the LMS and nudges the students' interest on getting the awareness. To accomplish this goal, a gauge version of the average-similarity visualization will be offered to the student (see Figure 3), offering the score estimation and with a see-more link. If the gauge caught the student's interest, then he will click on it and will access the A4Learning tool, where he will be able to select the simplified or the scatter versions of the visualization.

Figure 3: Gauge-based score estimation for its integration at the LMS.

4. Architecture of A4Learning and link to the LMS

The web based implementation of A4Learning allows its integration in heterogeneous learning environments, supporting cloud environments as well as other working contexts. In other words, the system is not part of any existing LMS or any other learning environment but instead it can be integrated in web based learning environments via widgets that offer a simplified version of the visualization and nudge the student to access A4Learning to view the complete information.

Furthermore, data collection is out of the scope of A4Learning, so it needs to be feed with activity logs representing the events generated by the students. This allows the analysis of web based scenarios as well as other learning environments whose activity can be tracked. For example, it could be feed by data collected from face-to-face scenarios that uses virtual appliances to monitor students' activity [6] or can support blended learning scenarios involving different spatial locations tracked with the help of NFC technology [16]. As different systems produce different data outputs, A4Learning uses MongoDB to store the information. Thus, the non-relational approach allows for an easy method to transform collected data from its original source into the format used by A4Learning.

The above leads to an architecture, depicted in Figure 4, that comprises the following modules:

- Data storage module. A no-SQL approach like MongoDB is used in order to store the students' activity logs. There subsystems are in this module:

  - o As these logs can be collected in various formats, a conversion module is required to translate the format and store the data in MongoDB. The conversion is case-specific (i.e. different learning tools will provide logs in different formats), so the transformation tool is extendable via plugins.
  - o Once the data has been properly transformed, it is stored in MongoDB and merged with the information already in the storage. As A4Learning can collect logs from different tools, the raw data storage

merges the input from these different sources.

o The module also stores a cached version of the analysis results. This means that the similarities are not calculated at the moment they are required. Instead, they are updated with a configurable frequency and stored back in the MongoDB database.

- Analysis module. This module reads the raw activity logs and calculates the similarity among students. Such processing is periodically triggered, so the visualization will be created with fresh information. The refresh frequency is initially set to "once a day", but can be configured to fit the needs of different scenarios.

- Visualization module. This module uses the cached data and creates the different visual representations of the similarity information, as described in the previous section. The produced visualizations can be used to feed the user interface, or can be integrated in an external tool (i.e. the LMS) via widget.

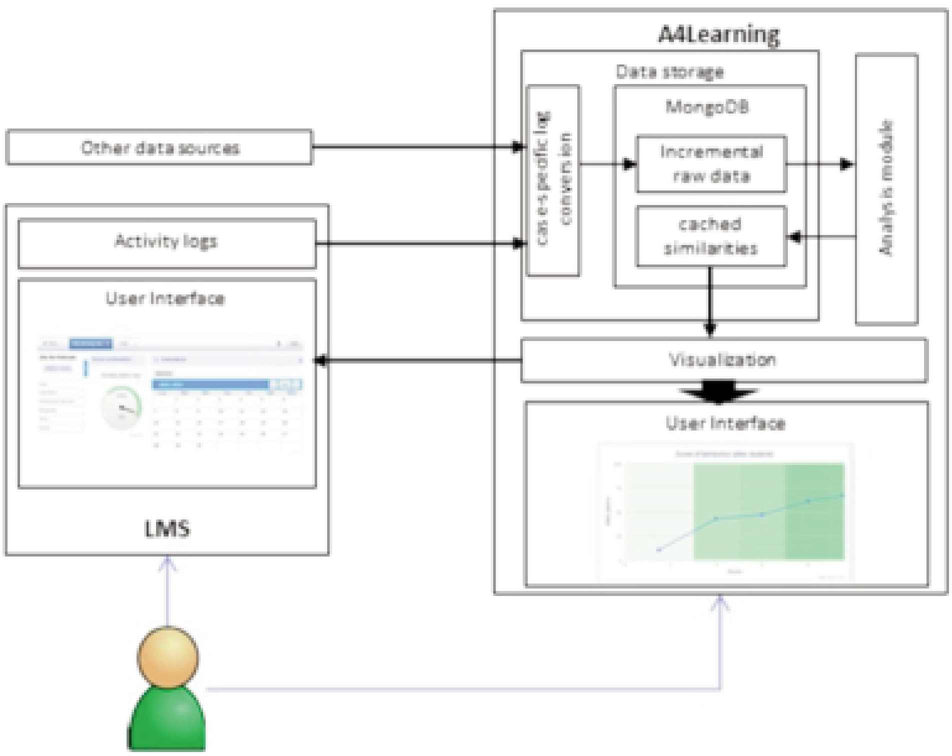

Figure 4: A4Learning modules, and connection to the LMS.

5. Conclusions and Future work

This article presents A4Learning, a tool that explores the idea of similarity between students and its relationship with the course achievements. Thus, a student can find similar students from previous courses and see the achievements of these students. This information is visually presented to the students, so they may find a rough estimation of their own achievements.

A4Learning has been implemented as a prototype that has served to evaluate its accuracy as estimation method and also to evaluate its usability by a group of experts. Within this first validation of the prototype, several possible improvements were identified, resulting in a simplified visualization that considers average values instead of showing all the information. The initially proposed scatter plot is maintained as "advanced view"

© ISSN: 2255-1514

and has been enhanced to provide more details on demand.

The web nature of A4Learning makes possible to integrate it in different learning platforms as a cloud service. To this extent, the learning platform is responsible to track the students' activity, which is be offered to A4Learning in the shape of activity logs. A4Learning analyses the information, calculates similarities, and represents them in relation to learners' achievements. Such visual representation is offered as a widget for its lightweight integration in the learning platform.

Next research efforts will focus on the validation of A4Learning in a real distance learning scenario. The first validation step will focus on usability, in order to increase the perceived usefulness and to increase the perceived self-efficacy of the students. This first validation step will result in the implementation of several adjustments both in functionality and interface. Then, a pilot program will validate the actual utility of A4Learning in the learning context.

Acknowledgements

This work is partially funded by UNIR Research (http://research.unir.net), Universidad Internacional de La Rioja (UNIR, http://www.unir.net), under the Research Support Strategy [2013-2015].

References

[1] M. Cocea, and W. Stephan. (2009). "Log file analysis for disengagement detection in e-Learning environments." User Modeling and User-Adapted Interaction 19, no. 4. pp. 341-385.
[2] J.P. Campbell, P.B. DeBlois, and D.G. Oblinger. (2007). "Academic analytics: A new tool for a new era." Educause Review 42, no. 4.
[3] M. Munoz-Organero, P. J. Munoz-Merino, and C. Delgado Kloos. (2010). "Student behavior and interaction patterns with an LMS as motivation predictors in E-learning settings." IEEE Transactions on Education, 53, no. 3. pp. 463-470.
[4] V.A. Romero-Zaldivar, A. Pardo, D. Burgos, and C. Delgado Kloos. "Monitoring student progress using virtual appliances: A case study." Computers & Education 58, no. 4. 2012. pp.1058-1067.
[5] M. Wang, J. Peng, B. Cheng, H. Zhou, and J. Liu. (2011). "Knowledge Visualization for Self-Regulated Learning." Educational Technology & Society 14, no. 3. pp. 28-42.
[6] S. Govaerts, K. Verbert, E. Duval, and A. Pardo. "The student activity meter for awareness and self-reflection." In CHI'12 Extended Abstracts on Human Factors in Computing Systems, ACM. pp. 869-884.
[7] J.L. Santos, K. Verbert, S. Govaerts, and E. Duval. (2012). "Addressing learner issues with StepUp!: an Evaluation." In Proceedings of the Third International Conference on Learning Analytics and Knowledge. ACM, 2013. pp. 14-22
[8] O. McGrath. (2010). "Data Mining User Activity in Free and Open Source Software (FOSS)/Open Learning Management Systems." International Journal of Open Source Software and Processes (IJOSSP) 2, no. 1. pp. 65-75.
[9] C. Romero, S. Ventura, and E. García. (2008). "Data mining in course management systems: Moodle case study and tutorial." Computers & Education 51, no. 1. pp. 368-384.
[10] G. Siemens, and R. Baker. (2012). "Learning analytics and educational data mining: Towards communication and collaboration." In Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, ACM. pp. 252-254.
[11] D. Keim, G. Andrienko, J.D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon. (2008). Visual analytics: Definition, process, and challenges. Springer Berlin Heidelberg.
[12] J.A. Larusson, and R. Alterman. (2009). "Visualizing student activity in a wiki-mediated co-blogging exercise." In Proceedings of the 27th international conference extended abstracts on Human factors in computing systems, ACM. pp. 4093-4098.
[13] R. Mazza, and V. Dimitrova. (2004). "Visualising student tracking data to support instructors in web-based distance education." In Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters, ACM. pp. 154-161.
[14] J. O'Donovan, B. Smyth, B. Gretarsson, S. Bostandjiev, and T. Höllerer. (2008). "PeerChooser: visual interactive recommendation." InProceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM. pp. 1085-1088.
[15] C. Carmean, and P. Mizzi. (2010). "The Case for Nudge Analytics." Educause Quarterly 33, no. 4.
[16] de-La-Fuente-Valentín, L., Pérez-Sanagustín, M., Hernández-Leo, D., Pardo, A., Blat, J., & Delgado Kloos, C. (2014). Technological support for the enactment of collaborative scripted learning activities across multiple spatial locations. Future Generation Computer Systems, 31, 223-237.

Cómo citar este artículo / How to cite this paper

De La Fuente, L. & Burgos, D. (2014).¿Lo estoy hacien bien? A4Learning como una herramienta de auto-conciencia de integración en los Sistemas de Gestión de Aprendizaje. Campus virtuales, 3(1), 32-40.

De La Fuente, L. & Burgos, D. (2014). Am I doing well? A4Learning as a self-awareness tool to integrate in Learning Management Systems. Campus virtuales, 3(1), 32-40.