Learning Analytics Dashboards for Professional Training - Challenges and Proposal

Mohamed Mouaici[12], Laurence Vignollet[1], Christine Galez[1], and Mael Etienne[2]

[1] Savoie Mont Blanc University, France
{mohamed.mouaici,laurence.vignollet,christine.galez}@univ-smb.fr
[2] Logipro company, France
{mohamed.mouaici, mael.etienne}@logipro.com

Abstract. Exploiting the large quantities of traces left by learners in Virtual Learning Environments (VLE) allows educators, learners and administrators to gain new insights into the learning process. Learning Analytics (LA) aims to leverage data collection, measurement, analysis and reporting data which can help users to improve the learning process. This paper presents the first results of the work we are conducting in a professional learning context to design an effective learning analytics dashboard. We show the particularities and explain the different challenges of our context that have led us to propose models to tackle it. We discuss how these models meet the requirements of our domain, and we finally give an example of indicators, measures and visualization built with educators to help them better understand the learner's behavior.

Keywords: Learning Analytics, Professional Training, Information Visualization, Measures, Indicators, Challenges, Models.

1 Introduction

Learning Analytics (LA) aims at exploiting the large amounts of data generated by the widespread use of on-line learning environments, such as Learning Management Systems (LMS) [1]. The Society for Learning Analytics Research defines LA as *"the measurement, collection, analysis, and reporting of learning data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs"*. Therefore, LA can be used to feed certain applications, such as Educational Data Mining (EDM) algorithms and Learning Analytics Dashboards (LAD) [2], which represent one of the most important areas of intervention of LA [2,3]. A LAD is considered as a container of indicators [3] calculated from many types of trace, e.g. resource use, time spent on platform, social interaction, assessment, manually reported data, artifacts produced [3,4]. These indicators can be categorized by type: learner-related, action-related, content-related, result-related, context-related and social-related indicators [3]. However, what the valuable indicators are remains a largely unresolved question that depends on several parameters, such as the context and the objective of the dashboard. In the literature, we can find some simple LAD with predefined indicators, such as Course Signals and StepUp! [2]. Other LAD are dynamic with customizable indicators, such as DDART [3]. In addition, certain LAD (such as Course Signals) use a prediction model to estimate and visualize the learning outcomes [2]. The prediction can also concern the engagement and at-risks learners. However, the studies focusing on these kinds of prediction model are still limited, as are the works on professional training that propose LAD solutions[3,5].

Hence our research project aims at proposing an LAD dedicated to a professional training context. Our research is conducted in collaboration between a company and an academic laboratory. The company is specialized in the development of IT solutions dedicated to professional learning and has its own LMS whereby training is delivered. The training provided concerns professionals, such as employees, entrepreneurs, craftspeople, etc. Moreover, the company's staff includes educators and instructional designers, in order to ensure the pertinence of the different steps related to professional training. We focus on the particularities of this domain (see Sect. 2) to propose relevant models that support educators in the creation of an effective LAD, by helping them to better understand the different situations occurring during the learning process and determine the learners' difficulties, in order to be able to assist them when necessary.

Based on this fact, our work addresses the following general research question: how to design an effective learning dashboard whilst taking into account the particularities of professional training?

2 Certain Particularities and Challenges of Our Domain

In our project we follow a user-centered design process [6]. So, we conducted the needs analysis study based on several interviews and surveys, including 12 users working for the company: 2 educators, 2 instructional designers, 4 software developers and 4 sales representatives. This study allowed us to identify the following particularities related to our domain of on-line professional training:

1. *Heterogeneous courses*: the company provides some trainings sessions that last only 14 hours. This kind of course is recurrent and aims at giving an introduction to the learner on a given concept. Other courses can last up to 700 hours, where the learner does the training over several months.
2. *The courses' duration is predefined and fixed:* in France, a professional training course may be funded by several players, such as "Pôle Emploi", which is an organization that funds the training of job-seekers. In addition, companies may pay for training for their employees. The price of the training is established according to its duration. Because of this reason, the duration is predefined, fixed and imposed on the learner by the funding organization.
3. *No possible dropout:* the funding of the professional training makes it mandatory to finish and does not give the learners the possibility of dropping out.

These particularities have a direct impact on the designing of a learning analytics dashboard solution. For example, the total time spent, which is considered as one of the important data sources in the literature [2], cannot always be relevant in our context. Indeed, all learners will spend the same total amount of time on training (because the course duration is predefined and fixed). So, building the analysis using this dimension (time spent) may be irrelevant and not representative of the learner's behavior. The no-dropout constraint forces the learners to finish their training even when they do not want to. This constraint can also have a negative impact on the engagement of learners because they feel obliged to finish the course. This situation needs to be identified by the educator in order to support learners to overcome it. Otherwise, the course's heterogeneity makes the task of identifying common relevant indicators very difficult. Indeed, certain relevant indicators in long courses may not be relevant for short ones. For example, "the learner connection frequency" may be a relevant indicator in a long course to indicate the continuous presence of learners in the platform. However, for a course that lasts only 14 hours, this indicator is not significant.

Section 3 presents the main models we propose in order to design an effective LAD that covers the particularities described above.

3 Proposed Models

In addition to the particularities of our study domain (see Sect. 2), with the end-users (educators) we identified a need to focus on particular measures to facilitate the analysis of a specific situation related to professional learning. The need to focus on three measures was identified: learner's progress, learner's engagement and at-risk learners. In fact, in order to perform a focused analysis, the educator can, for example, ask to see only indicators representing the engagement of learners. So, defining a measure (engagement for example) consists in grouping a set of indicators in the same visualization page of the dashboard. However, which indicators are likely to represent given measures remains a largely unresolved issue [7]. Furthermore, in our study we observed certain differences in terms of indicators to define a particular measure depending on educator's experience (see Table 1). The difference also concerns the visual components used to display an indicator. Indeed, with the variety of the possible visual components to display an indicator, educators have preferences for those which are more familiar to them. Moreover, the educators involved in our study showed an interest in being assisted by the system to determine the at-risk learners. This functionality would allow them to focus their help on learners who need it most. With the variety of needs and the particularities identified, we propose a customizable solution based on models to create an effective LAD dedicated to our professional learning context:

1. *Data model:* it allows the management and formatting of learning traces in a specific format before storing them. Our model uses xAPI which is an event-centered specification that makes it possible to collect a wide range of learning traces (also called learning experiences or statements)[8]. In addition, using xAPI specification facilitates the extensibility and interoperability of learning traces collection architectures [8]. In our model, we define the

relevant events (user actions to be tracked) related to our context and the process of creating statements based on simplified xAPI specifications[8]. The created statements will be stored in a Learning Record Store (LRS). This model is materialized by trackers, that retrieve traces from learning environments (from an LMS and a Forum in our context), transform them into xAPI statements and insert them in the LRS. If the retrieved traces are already in xAPI format, they will be directly inserted in the LRS.

2. *Filtering and aggregation model:* it enables the calculation of a predefined set of indicators by reading the xAPI statements stored in the LRS. In this model, we define a data processing procedure that includes data cleaning, data transformation, indicator calculation and measure definition. For example, from a learner's logs, we can calculate the learner's connection frequency and aggregate it per week or per month. In addition, this model will determine which measures include this indicator for which educator. As we provide educators with the possibility of creating their own indicators and measures, this model is responsible for managing indicators and measures defined by each educator.
3. *Visualization model:* it aims at determining for each indicator an adequate visualization component (box plot, bubble chart, pie chart, table, etc.). Based on common features between indicators and on supported visualization components, this model helps to identify the appropriate visualizations that can be used to display an indicator. In addition, it allows the definition of the granularity and the aggregation level (the quantity) of data to be presented per: learner, group of learners, day, week, etc. By default, for each indicator we propose an adequate visualization based on this model, and then, users can modify the visualization according to their preferences. Based on these features, the system supports them in this choice by providing them with the list of the visual components likely to display the indicator.
4. *Prediction model:* it analyses data to estimate the probability that a learner will be at risk. The major challenge in this model is to be able to predict risks in short courses (that last 14 hours for example), where defining the learner's behavior model can be very difficult. The prediction will be translated by notifications sent to the educator's LAD to inform them about the at-risk learners. Otherwise, as dropout does not exist in our context, the risk concerns:
   - *Risk of failure:* the learner fails the training course if the score obtained is less than 70 %.
   - *Psycho-social risks:* they are associated with the overwork that an employee can experience when working and doing the training at the same time; doing the training in the evenings and on weekends. This situation impacts directly on the performance of the learner and can cause the learner's failure [9].

The main requirements of the proposed models are to tackle the particularities identified in our context and to facilitate the adaptation of our solution to additional constraints. Indeed, the proposed solution provides users with a framework allowing them to: construct their own indicators and measures, which can help them to adapt their LAD to different scenarios; choose adequate visualization to let them get a relevant familiar representation and which respects their knowledge visualization level; receive real-time notifications to facilitate their intervention to help at-risk learners. In addition, the proposed solution needs to be easy, allowing educators to use it to create indicators and define measures without the need for advanced statistical and computer skills.

Figure 1 depicts the components of the proposed solution related to the above-mentioned models and shows the interactions between them. Each number indicated in Fig. 1 corresponds to the model associated with this number in the description.

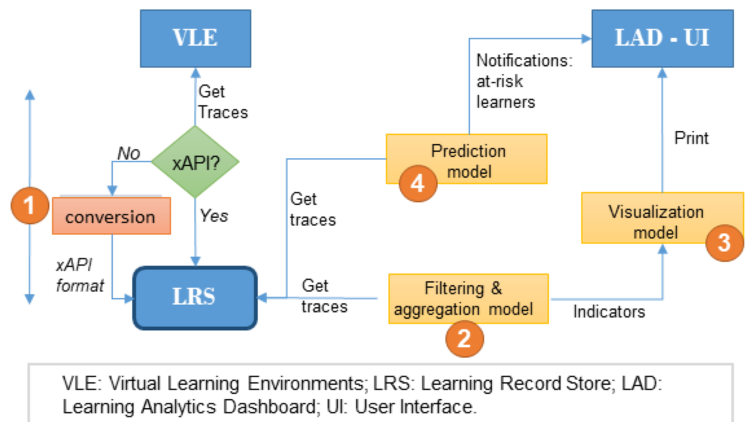

Fig. 1. Proposed models and their interactions

Table 1 gives an example of measures with their relevant indicators as well as the adequate visual components defined by two educators involved in our study.

4 Conclusion and Perspectives

In this paper we have discussed certain particularities and challenges related to our professional training context. These peculiarities lead to different needs for educators in terms of measures, indicators and visualization components. To tackle these requirements, we have proposed a Learning Analytics Dashboard (LAD) solution based on models to allow its adaptation to different uses and constraints. Thus, the proposed models provide users with a predefined set of indicators displayed on adequate visualization. Then, users can define measures by selecting and grouping indicators. Moreover, the users can create their own indicators by setting (defining the traces that will be used to construct the desired indicator) the data filtering model. The visualization model allows users to determine which adequate visual components can be used to display an indicator,

Table 1. Examples of indicators, measures and adequate visual components

Learning traces	Indicators		Indicator Visualizations		Measures
	Educator 1	Educator 2	Educator 1	Educator 2	
Connection date, content visualization, mark obtained, comments, forum posts	Connection frequency per day, attempts/quiz, content rate visualization.	Number of forum posts/ comments, attempts/quiz.	Line graph, bar chart.	Network graph, bar chart.	Engagement
	Activity completion rate, degree of skill acquisition	Completion rate, learning path.	Gauge chart.	Gauge chart, network graph.	Progress

which provide them with the possibility of choosing a more familiar visualization component that respects their knowledge level. The prediction model helps educators to identify the at-risk learners by sending them notifications on the LAD. The challenge is to train this model to predict at-risk learners in short courses. The other challenge will be to study the impact of our solution on the engagement and the success of learners in our domain.

References

1. Elias, T.: Learning analytics. In: Learning (2011)
2. Verbert, K., Govaerts, S., Duval, E., Santos, J., Van Assche, F., Parra, G., Klerkx, J.: Learning dashboards: an overview and future research opportunities. In: Personal and Ubiquitous Computing, Vol. 18(6), pp. 1499–1514. Springer (2014)
3. Schwendimann, B., Rodriguez-Triana, M., Vozniuk, A., Prieto, L., Boroujeni, M., Holzer, A., Gillet, D., Dillenbourg, P.: Perceiving learning at a glance: A systematic literature review of learning dashboard research In: IEEE Transactions on Learning Technologies pp. 30–41 IEEE (2017)
4. Bodily, R., Verbert, K.: Review of research on student-facing learning analytics dashboards and educational recommender systems. In: IEEE Transactions on Learning Technologies, Vol. 10(4), pp. 405–418. IEEE (2017)
5. Littlejohn, A.: Learning and Work: Professional Learning Analytics. Society for Learning Analytics Research (2017)
6. Abras, C., Maloney-Krichmar, D., Preece, J.: User-centered design. In: Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications, Vol. 37(4), pp. 445–456. Citeseer (2004)
7. Donaldson, L., Matthews, A., Walsh, A., Brugha, R., Manda-Taylor, L., Mwapasa, V., Byrne, E.: Collaborative Tools to Enhance Engagement in a Blended Learning Masters Programme. In: AISHE-J: The All Ireland Journal of Teaching and Learning in Higher Education (2017). Vol. 9(1).
8. Santos, J., Verbert, K., Klerkx, J., Duval, E., Charleer, S., Ternier, S.: Tracking data in open learning environments. In: Journal of Universal Computer Science, Vol. 21(7), pp. 976–996. Springer-Verlag (2015)

9. Marc, J.: Le recours aux formations à distance (e. learning) dans la formation professionnelle des salariés. Présentation, influence sur les acteurs et éléments de vigilance, INRS (2014)