HAL open science

METALRS: towards effective Learning Analytics through a hybrid data collection approach for the french lower secondary education system

Laura Infante-Blanco, Azim Roussanaly, Anne Boyer

To cite this version:

Laura Infante-Blanco, Azim Roussanaly, Anne Boyer. METALRS: towards effective Learning Analytics through a hybrid data collection approach for the french lower secondary education system. 2nd Annual Learning & Student Analytics Conference, Oct 2018, Amsterdam, France. hal-02469611

HAL Id: hal-02469611

https://inria.hal.science/hal-02469611v1

Submitted on 6 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

METALRS: towards effective Learning Analytics through a hybrid data collection approach for the french lower secondary education system

Laura Infante-Blanco, Azim Roussanaly, Anne Boyer

Université de Lorraine

Indicate your track:

1. Academic research: comprehensive evaluations of recent innovations in learning and student analytics approaches.

Purpose

Data collection is one of the most critical and time-consuming tasks when dealing with Learning Analytics (LA) system design. This problem has been addressed in the context of METAL [6] (Modèles Et Traces au service de l'apprentissage de Langues), a project whose aim is to improve the learning experience of lower secondary education students of the Lorraine region in France. Our collection approach follows a hybrid methodology between a top-down Data Mining (DM) approach based on the CRISP-DM process [1] and a bottom-up analysis starting from the international e-learning specifications [9] Experience API (xAPI) [4] and OneRoster [5]. As a result of this hybrid methodology, we have developed METALRS [8], an extended Learning Record Store (LRS) that combines standard student learning traces with academic data and other more intrinsic student information in a multidimensional model.

Design

The CRISP-DM methodology defines a process model that structures and describes the different phases involved in a DM project, namely: (1) *business understanding*, (2) *data understanding*, (3) *data preparation*, (4) *modeling*, (5) *evaluation*, and (6) *deployment* [2]. During the *business understanding* phase (1), we have conducted a deep study of the regional lower secondary education system in order to identify its major actors and stakeholders, as well as the main LA goals to be achieved. In that context, the first axis of the METAL project focused on the construction of teacher dashboards and student evolution panels able to reflect not only *past* and *present* activity indicators but also predictions of *future* learning outcomes and remediation actions. During the *data understanding* phase (2), we identified the most

relevant educational activities and processes, the links between them, and the data resulting from them. This data is distributed between many academic agents, i.e., the regional education academy, each one of the four county councils, each high school, and also in open and private online resource systems. The *data preparation* phase (3) implied the coordination of the various stakeholders as well as the process of mapping the data into a single naming convention that is applied directly into the source. At this stage of the CRISP-DM process, we have built a conceptual model in agreement with the xAPI specification (to represent students interactions with tools in the learning environment [3]) and an extended OneRoster model (to acquire present and historical academic data, demographics, and personal student information), according to the modeling needs predefined during the *business understanding* phase (1). The *data modeling* phase (4) uses the previous conceptual model to ease the selection of multiple DM techniques in order to address the numerous LA challenges [10] that METAL aims to resolve. For example, using different sequence pattern mining techniques over activity and learning traces in order to generate online learning resources recommendations to students and enhance their training in their weakest areas.

Results

DM processes like CRISP-DM constitute an effective approach for learning and student data collection that can be combined with the use of international e-learning standards to ensure adaptability. In the context of the METAL project, CRISP-DM guided us through the discovery of the french national education ecosystem from a data point of view [7].

METALRS [8] constitutes an intermediate product of the application of the CRISP-DM methodology to the METAL project. Our generic conceptual model and its implementation serve as a high-level framework to produce several sub-models. It is important to note that our approach involves the collection of learning and student data from many different sources. Therefore, some of this input data will require a transformation and translation process from text files to generate either xAPI or oneRoster compliant representations. According to the type of data, the collection will be carried out at different time intervals, e.g., learning and

activity traces might be sent in real time while personal student information might be collected at the beginning of every semester. The collected data is then processed by METALRS services and stored in a database. Afterwards, the obtained data is ready for further analysis, visualization, and also to be used as input for learning sub-models in charge of predicting and remediating student learning outcomes. Figure 1 describes the overall architecture of METALRS as well as the involved external components.

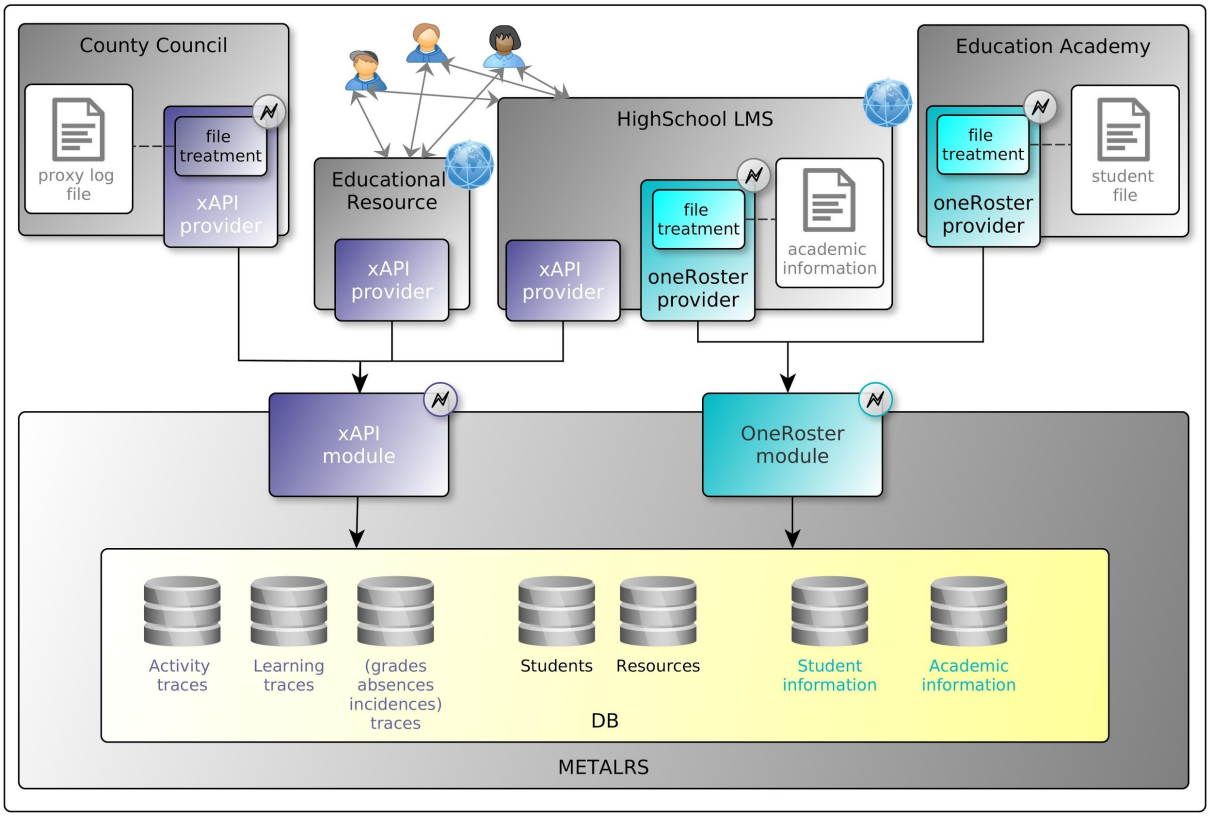

Figure 1: Overall METALRS architecture

Implications

The french lower secondary education system is a singular environment because of the diversity of, online and offline, educational and learning management systems. The establishment of an approach for centralization and unification of data constitutes the first step towards a smart integrated educational system and a fundamental pillar for effective LA.

Acknowledgments

This work has been carried out within PIA E-Fran METAL which receives funding of Programme d'Investissements d'Avenir, France.

Resources

1. Mariscal G., Marbán Ó., Fernández C. (2010). A survey of data mining and knowledge discovery process models and methodologies. In *The Knowledge Engineering Review*, 25(2), (pp. 137–166). Cambridge University Press.

2. Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C. Wirth, R. (2000). CRISP-DM 1.0 Step-by-Step Data Mining Guide. Technical report, CRISP-DM

3. Kevan, J. M., Ryan, P. R. (2016, Apr). Experience API: Flexible, Decentralized and Activity-Centric Data Collection. Technology, Knowledge and Learning, 21(1), (pp. 1-7).

4. Experience API. https://github.com/adlnet/xAPI-Spec/blob/master/xAPI.md. Accessed : 2018-07-24

5. IMS OneRoster Specification Version1.1. https://www.imsglobal.org/oneroster-v11-final-specification Accessed : 2018-07-24

6. Projet eFRAN METAL : http://metal.loria.fr. Accessed : 2018-07-24

7. Roussanaly A., Toulotte T., Infante-Blanco L., Boyer A., Brun A., Bonnin G. (2017) Projet METAL: Plan de collecte de données. Report, Université de Lorraine (Nancy); CNRS. (in French)

8. Roussanaly A., Infante-Blanco L. (2018) Projet METAL: Fondements pour la conception d'un entrepôt de données d'apprentissage. Report, Université de Lorraine (Nancy); CNRS. (in French)

9. del Blanco Á. , Serrano Á., Freire M. , Martínez-Ortiz I. and Fernández-Manjón B. (2013). In E-Learning standards and learning analytics. Can data collection be improved by using standard data models? IEEE Global Engineering Education Conference (EDUCON) (pp. 1255-1261). IEEE.

10. Siemens, G., & Baker, R. S. (2012, April). Learning analytics and educational data mining: towards communication and collaboration. In proceedings of the 2nd international conference on learning analytics and knowledge (pp. 252 - 254). ACM.