Requirement-driven Architecture for Learning Analytics: The Case of AT41 Project

Nesrine Zibani, Ladjel Bellatreche, Hassina El Kechai, Sébastien Iksal

Requirement-driven Architecture for Learning Analytics: The Case of AT41 Project.

N. Zibani*, H. El Kechai, L, Bellatreche and S. Iksal

Abstract: Digital technologies are becoming more present and essential in all sectors of our life. In education, the intensive usage of digital learning devices contributes to generating a large amount of trace data from digital learning activities. Intelligent exploitation of these traces represents a valuable asset for both device producers (to improve the design of the devices) and consumers (learners and teachers). In this paper, we first share our vision for better exploitation by teachers, of traces from middle schoolers' digital activities generated by their use of tools and digital learning services during different classes. This vision is a part of the AT41 project funded by the French Ministry of Education. This exploitation has to meet the requirements of the different teachers. Conducting such a project is not an easy task, because it has to consider the following issues: (i) the lack of comprehensive and clear methodology to design and exploit these traces, (ii) heterogeneity of teacher requirements that complicates their elicitation and analysis, and (iii) the diversity of trace sources. Secondly, we propose a requirement-driven architecture for Learning Analytics composed of a well-identified life cycle. This architecture is augmented by learner traces. It offers a repository storing both teacher requirements and traces to facilitate the Learning Analytics in generating relevant and valuable indicators.

Key words: Dashboard, Learning Analytics, Trace Exploitation, Requirements.

· N. Zibani. Zibani is with LIAS, ISAE-ENSMA, Poitiers and 86000, France. E-mail: nesrine.zibani@ensma.fr
· H. El Kechai. El Keachai with Techne, Poitiers University, Poitiers and 86000, France. E-mail: hassina.el.kechai@univ-poitiers.fr
· L. Bellatreche. Bellatreche is with LIAS, ISAE-ENSMA, Poitiers and 86000, France. E-mail: ladjel.bellatreche@ensma.fr
· S. Iksal. Iksal is with LIUM, University of Le Mans, Laval and 53000, France. E-mail: sebastien.iksal@univ-lemans.fr
* To whom correspondence should be addressed.
Manuscript received: year-month-day; revised: year-month-day; accepted: year-month-day

1 Introduction

In information and communication, the technological environment is continuously changing and consequently impacting our lifestyles. Education with its Eco-system is one of the most important pillars in our life. The school is a noble institution that trains and educates future citizens. Thus in a digital era, schools spend a lot of money to move from traditional practices to digital practices. A digital practice is a set of frequent and usual instrumented thematic actions, constructed in the interaction with an object or the same environment aiming at a certain efficiency. In this context, the French Ministry of Education deployed a Digital Plan for Education from 2015 till the end of 2017, where 1668 middle schools were equipped with various digital equipment [14]. Following this plan, the Directorate for Digital in Education in its Numéri'Lab incubation mission, financed various projects, in priority the ones dealing with Learning Analytics' topics.

Learning Analytics is the measurement, the collection, the analysis, and the reporting of data about learners and their contexts, in order to better understand and optimize learning and the environments in which it occurs [11]. Our work is part of an incubator project of the Orleans-Tours academy entitled AT41 [1]. This project aims mainly at analyzing the digital practices of Loir-et-Cher's public middle school pupils that are equipped in class with Sqool tablets issuing digital traces. A Sqool[2] tablet is a French solution with an e-educational system dedicated, adapted to each cycle, and offering three spaces associated to the learner, the teacher, and the school headmaster. The explosion of traces issued from these devices represents an asset for teachers and school managers for building smart educational decision support systems to enhance the teaching activities.

The exploitation of traces has to meet the requirements of the different teachers. These participating teachers are associated to different persona profiles created based on teachers' requirements. Crucial issues have to be addressed when conducting our project: (i) the absence of comprehensive methodology to design and exploit these traces adapted to the context of French middle schools; (ii) the elicitation and analysis of teacher requirements are not an easy task since requirements are heterogeneous. More precisely, they are given by various teachers using different formalisms, multimodal formats (text, audio, video, drawing) and syntax; and (iii) the diversity of the profile of learners.

This situation motivates us in proposing a requirement-driven architecture augmented by traces for Learning Analytics. This joint consideration of requirements and data is a common approach in constructing valuable Decision Support Systems [2]. It is a common approach composed of a well-identified life cycle that includes the following phases: (1) a middle school composed of teachers and pupils, respectively issuing requirements and traces, (2) the requirement elicitation and analysis that requires the carrying out of workshops in order to collect relevant needs of teachers involved in diverse teachings. The process of analyzing teacher requirements will allow us to understand different activities, their performance indicators, and their evaluation. (3) The trace pre-processing, (4) the storage repository, in charge of storing both teacher requirements and learner traces, (5) exploitation by different tools such as visualization, querying, data mining, recommendation, and so on and (6) the capitalization of indicators for the reuse purpose. It should be noticed that the first five steps of our architecture are quite similar to existing Big Data Analytics solutions proposed in various domains such as Smart Cities [1]. The majority of these solutions ignore the capitalization phase. These different generic phases of our architecture will be deeply detailed in this paper, by instantiating them use the school of *Pierre de Ronsard, located at Mer, France*[3].

The remainder of this paper is organized as follows: Section 2 presents different research issues related to our project. Section 3 gives a state of the art on Learning Analytics, dashboards, and indicators as well as our positioning within existing works. In Section 4, we describe our architecture by detailing its different components illustrating by examples. Section 5 concludes our paper. .

2 Problem and Challenges

Several works are interested in the analysis of digital traces of the learning activity. As far as we are concerned, we are interested in digital activity that supports the learning activity when it is realized with different digital technologies and devices. The analysis of digital practices is a good indicator of the progress of that activity.

[1] From the Appropriation of digital tools to the Transformation of teaching practices in the Loir-et-Cher department (41)

[2] http://www.sqool.fr

Therefore, our objective is to identify and develop one or more learning dashboard structures restoring pupils' digital activities meeting the teacher's requirements. These structures shall at the same time ensure reuse, sharing, and evolution from one middle school context to another, in particular, the capitalized indicators. To solve such problem, we attempt to answer the following research questions:

- What are the teacher requirements in terms of restoring juvenile digital practices via dashboards? Answering this question allows us to categorize different digital practices.

- What are the indicators that emerge from these teacher requirements? What are the appropriate metrics for evaluating the quality of indicators? In [12], an indicator is defined in general as "an evaluation or information tool (device, instrument, quantity) which should serve as an aid to decision-making". Answering this second question leads us to define and structure the indicators per category of digital practices. The evaluation metrics of these indicators is performed using traces. This data has to be cleaned, transformed, and stored in a given repository in order to generate relevant indicator and their visualization in dashboards.

- How these indicators are capitalized, enriched, and reused in the analysis processes of juvenile digital practices? This question is fundamental since it implies the definition of capitalization approach augmenting the reuse of the identified indicators in other middle schools than those that have contributed to our project. Usually, data-driven decision support systems ignore the capitalization phase.

3 Related work

This section reviews major studies related to different components involved in Learning Analytics: Requirements, Activity, Data, and Indicators (see Fig. 1).

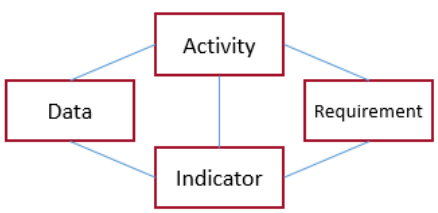

Fig. 1 Requirement-driven Architecture for LA concepts

Requirement gathering is a crucial step in Learning Analytics. In Hubble project, [12] identified the teacher requirements by interviewing them analyzing to personalize the learning of their 3rd cycle students in a digital environment at school and propose a model defining learning indicators. [19] aimed at examining how digital transformation affects universities and students. Semi-structured interviews were conducted and the analysis of the transcribed interviews was based on Mayring's qualitative content analysis [13]. In [6] researchers developed Learning Analytics and Educational Data Mining meeting the needs of students and teachers in pre-tertiary education. Requirements were gathered using specific groups, workshops, interviews, literature review, and research experience. They also proposed a new methodology for needs analysis for the learning analytics system in pre-tertiary education and claim that it becomes imperative to develop feasible needs analysis methodology to ensure that main groups of users get what they need and require.

The analysis of the digital activity by exploiting learning's digital traces has been widely studied in the literature in order to enhance teaching activities. The Living Cloud project[4] is an example of these studies. Also the work of [15] made it possible to conduct an analysis of the circulation of juvenile digital practices using their traces and proved the existence of a correlation between commitment, behavior, and academic performance based on the activity system model proposed by Engeström [8].

The Hubble project [5] aimed at creating a national observatory for the construction, development, and sharing of massive data analyzing's processes issued from digital traces generated in e-learning environments for several actors (such as teachers, designers, administrators or even politicians). The MULCE project [16] aimed at collecting and sharing traces and propose the diffusion of a learning data corpus called LETEC. Many are the tools and methods from French research concerning collection, storage, and trace analysis as DDART, Lab4ce, Tactiléo, Kidlogger, TraceMe, TRAVIS[6].

Several research efforts conducted to design dashboards. The authors of [17] present a systematic review of the literature on learning dashboards in Learning Analytics and Educational Data Mining fields.

[3] http://clg-ronsard-mer.tice.ac-orleans-tours.fr/php5/

[4] https://techne.labo.univ-poitiers.fr/living-cloud/

[5] http://hubblelearn.imag.fr

[6] GTnum2: http://techne.labo.univ-poitiers.fr/gtnum2/

They claim that learning dashboards are gaining popularity due to the increased use of educational technologies, such as Learning Management Systems (LMS) or Open and Massive Online Courses (MOOC) [17]. The work of [5] which is a part of Hubble focused on the identification of learning dashboard structures, where different models were proposed to describe users, and the context. An automatic process of generating learning dashboards based on the engineering knowledge approach [3] design using xAPI[7] and meeting the needs of primary and secondary school teachers in terms of learning activity's indicators. Recent studies were focused on the capitalization of indicators [10] to increase their reuse.

4 Our architecture

The architecture we implement is organized over four phases. First, two phases realized in parallel: A step of workshops to collect teachers' needs then elicit and analyze them to also identify digital practice categories required by teachers in order to produce instruments to be used to design and implement both the dashboards desired by teachers but also the approach to capitalize on indicators. The other one consists of pre-processing of the pupils' digital traces-data collected using tracing tools. Cleaning, transformation, and structuration of these traces is necessary in order to be appropriately stored. In the third phase, we store all the data already collected and structured in both previous phases. Finally, we pass in the last phase to the data processing stage where the data stored are exploited and used in different Learning Analytics formats such as dashboards or recommendation systems based on teachers' needs we developed the tools for.

Our Requirement-driven Architecture for Learning Analytics is presented in the Fig. 2

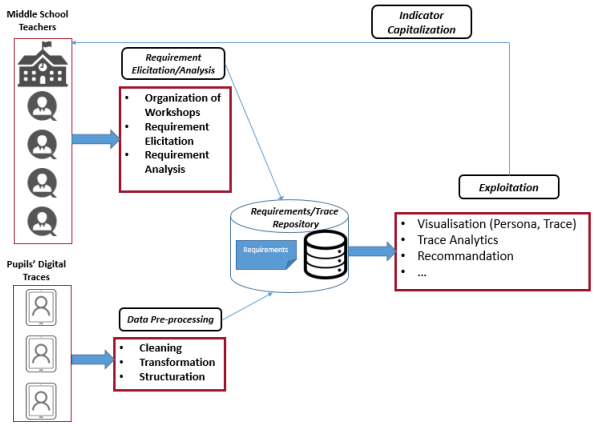

Fig. 2 Requirement-driven Architecture for LA

[7] http://experienceapi.com/overview

The approach phases illustrated by our implementation in AT41 project (Pierre De Ronsard middle school of Mer) case are detailed below:

4.1 Requirement Elicitation and analysis

This phase requires significant implementation efforts. On the one hand, teachers are not always aware of their needs, on the other hand, analysts do not necessarily have the same knowledge in education as teachers.

4.1.1 Workshops

As part of the first phase, participatory design workshops are organized in collaboration with pilot middle school teachers participating in the AT41 project. The objective of these workshops is to collect their relevant needs for reporting on their pupils' digital activity allowing teachers to express their needs and identify the information they wish to view on the dashboards.

Eight teachers were contacted by the school principal and accepted to participate in this workshop, which represents 30% of the college's teaching staff (5 teachers and 3 teachers). Several teaching subjects are represented: documentation, mathematics, English, technology, geography, history, and sport.

To realize this work, we organized 2 workshops for each middle school with an average duration of 3 hours. The two workshops take place in two different dates (see Fig. 3). These methodology phases are presented below:

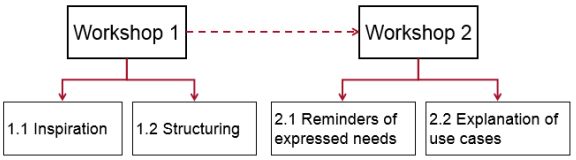

Fig. 3 Workshops general outlines

First phase of workshop 1 (Inspiration):

Individual work is initiated so that each teacher produces their own needs, then in groups teachers work together in order to identify common needs and then prioritize them according to their importance for them.

Second phase of the workshop 1 (Structuring):

In a group, teachers use a design cards kit [9] in order to categorize and structure their expressed needs in the first phase.

In this second workshop, the groups are formed based on the teachers' interest to the same needs and the same thematic categories of digital practices that they expressed in workshop 1. The second workshop also takes place in two phases:

First phase of workshop 2 (Reminders of needs expressed):

The needs expressed during the first workshop are structured according to two large thematic categories, one of the digital pupils' practices and another concerning teaching pedagogical practices. They are presented again to teachers in the new structured format.

Second phase of workshop 2 (Explanation of use cases):

This phase consists of each teacher individually to choose from the list of needs expressed and categorized previously the most relevant ones in order to describe use cases that make sense for their teaching practice. Teachers who share the same interest for the same needs gather to sketch their first projections in terms of content (indicators and objectives) and visualizations of dashboards.

During these workshop stages, we collected in addition to the needs expressed various other data like videos, audios, photos and also notes taken by teachers and researchers, a questionnaire has also been distributed to the participating teachers in order to understand their profiles, their digital practices, their uses of the Sqool tablets as well as their (potential) previous uses of dashboards.

4.1.2 Requirement Analysis

During the workshops, we collected a large amount of data in different formats. Most of this data can not be directly stored and exploited. Therefore, in this phase, we analyze and structure the data to bring them to such a state that they can be easily stored and interpreted.

4.1.3 Analysis methodology

Users' needs analysis is one of the main issues in the Learning Analytics context [6]. Within the AT41 project, after realizing the workshops, we developed the methodology for needs analysis for Learning Analytics-based on pupils' digital practices, we analyzed everything produced or collected from the workshops over 5 phases (see Fig. 4)

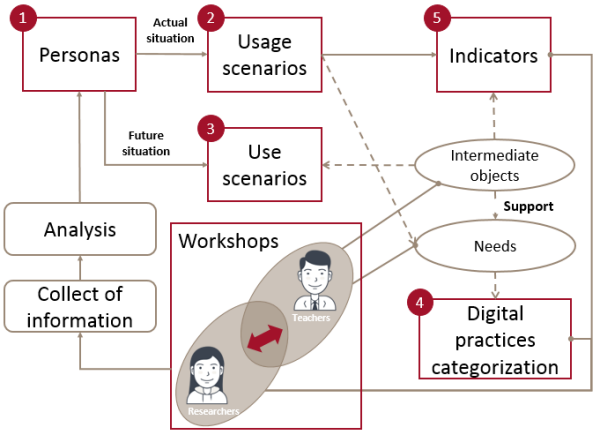

Fig. 4 Analysis process

(1) During the first phase, we collected information to create our target users personas [4]. We collected a set of information, in particular from questionnaires that we distributed to teachers before the beginning of the workshops. This allowed us to create user groups, structure them, and thus develop the persona for whom we will design the dashboards (see Fig. 5a)

(2) For each persona, we have modeled usage scenarios. Next, we identified situations-problem for which the dashboards could provide an answer. This usage scenario also describes the opportunities that dashboards could offer to respond to the problems posed (see Fig. 5b)

Fig. 5a Persona

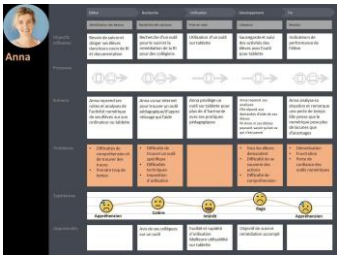

Fig. 5b Scenario

(3) The third phase involves concretely modeling how teachers project themselves in using the future dashboards in terms of human-machine interaction in use scenarios format. We proceeded to analyze the Intermediate Objects [18] produced for each need during the workshops. Thanks to these use scenarios, teachers describe anticipated actions interacting with the future tool, which we then modeled with use case diagrams.

(4) In a 4th step, we associated the needs expressed to thematic categories of digital practices.

(5) Finally, from the different productions and our analysis, we started transcribing projections in terms of visualizations produced by the teachers and develop the indicators.

The first indicators that emerged from the workshops are classified into four categories. We give here a non-exhaustive list:

- **View indicators**: to view statistics about the sites and files visited, the time spent on a site, the percentage of each site visited over the all or pupils state in front of the tablet (active, not active)
- **Information retrieval indicators**: analysis of all pupils searches: for example number of a successful search, search status, search keywords, or search duration

- **Monitoring indicators:** monitoring of search progress over time with the history recording
- **Comprehension indicators:** relate to autonomy, profiles of information searcher and digital behavior, the relevance of the keywords typed during a search, the methodology, and the strategy used for the information retrieval

In terms of visualization modes, the first projections suggest the need for simple visualizations considered effective by teachers: tables, monitoring trees, flow diagrams, word clouds, hourglass sand timer … a notification system left to the appreciation of the pupils in case of difficulty for more targeted remediation is also expressed. The teachers insist to have two interfaces, one dedicated to the teachers and the other one to their pupils with some common parts shared between them, which will serve as awareness tools [7] to bring pupils to operate a reflexive approach on their own digital practices in interaction with the teacher.

4.2 Data pre-processing

In order to track pupils' tablets, with the collaboration of the DSI (Department of Information Systems) of the Loir-Et-Cher department, a "Kidlogger" tracking solution that collects basic traces relating to the pupils' digital activity on Sqool tablets is installed. This tool was insufficient in terms of needed traces, so we complete the missing needed-data using "DPA" solution locally developed for a similar project. We claim that an easy data-exploitation process and good results depend, among others, on the quality of the data used. Therefore, before integrating traces in our repository, we clean, transform, and structure them.

4.3 Requirements/Trace Repository

In this phase, we aim at storing the data already treated in the previous phases (workshops and tracking). One of the most important data that we store is the indicators expressed by teachers and constructed during the previous phases. This model can be completed by realizing other workshops if other information was required and is not mentioned in the current model.

Indicator model: In order to develop the indicators in a complete and rigorous way, we decided to describe and model an indicator inspiring by the OVAR method[8] and the Unesco[9] which describes many indicators for education. We have created an indicator model as presented below (see Fig. 6)

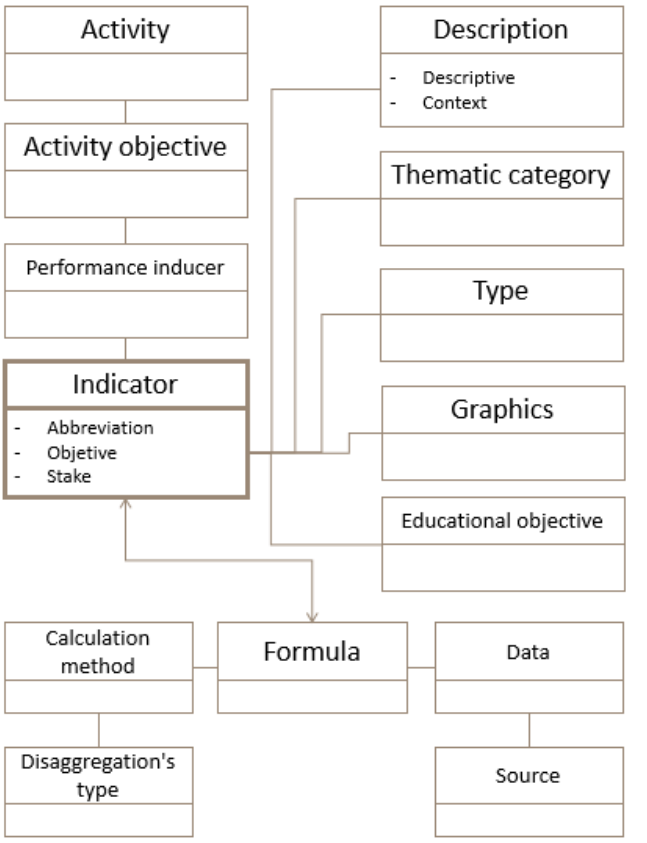

Fig. 6 Indicator model

We describe each concept of this model in Table 1 below:

Table 1 Model's concepts.

Concept	Description
Activity	The name of the activity to be treated or analyzed using a set of indicators
Activity objective	The goal we want to achieve by this activity
Performance inducer	Actions that induce to have a good performance and reach the objective
Indicator	The element to calculate, track and display
Abbreviation	Abbreviation of the indicator name
Graphics Visualization	The mode of the indicator on the dashboard adapted to the type of indicator
Description	A description and documentation of the indicator
Objective	The target of the present indicator
Educational objective	The objective motivated by the educational situation
Stake	Explanation of the impact of this indicator on the user
Data	The necessary data for the construction or calculation of the indicator

[8] Objectives Action Variables Responsibility; consistent French approach to designing a dashboard system.

[9] l'Institut de statistique de l'UNESCO. 2009. Indicateurs De L'éducation: Directives Techniques. (2009).

Source	Sources where we can recover data
Calculation method	The steps or approach to follow to calculate the indicator
Formula	Mathematical formula to calculate the indicator
Disaggregation ’s type	The different levels of calculation of the indicator

We instantiate the model using the AT41 project indicator example. For the persona Anna who represents teachers needing to restore the information retrieval strategies of her students, the Activity consists in perceiving the student's information retrieval strategy. Performance inducer: restoring the temporality of search. Indicators to identify sources of information and their relevance: domain names of the sites visited, extensions. Indicators: duration of site consultation, the total duration of information retrieval. Thematic category of digital practices: information retrieval. Type of indicators: monitoring. Educational objectives: Teaching pupils to select information sources taking into account domain extensions, diversify information sources, and make a search in a limited time. Type of disaggregation of these indicators: per student.

4.4 Exploitation

An important amount of data is stored in our repository. For our AT41 project, this mine of data allows us to developed relevant Learning Analytics in generating and capitalizing the required digital practices. The indicators displayed on the dashboard adapt to each persona requirements.

5 Conclusion

In this paper, we present a clear and comprehensive architecture for Learning Analytics. In addition to the fundamental phases dedicated to data-driven analytics solutions, our architecture includes an important phase in the education representing the **capitalization of indicators** generated by the process of Learning Analytics. This capitalization of indicators augments the reuse, share, and evolution potential to satisfy a large number of schools. All steps of our architecture have been implemented by considering the middle school of **Pierre de Ronsard, located at Mer, France**. The particularity of our architecture is that it is based on teacher requirements and pupils' traces generated during their digital learning activities. Before proposing this architecture, we identified different issues that have to be handled in middle schools in terms of eliciting and analyzing requirements, cleaning traces, and understanding different activities offered by each teacher profile. The fieldwork allowed us to identify different requirements and activities and contributed to generating indicators.
Currently, we are reproducing our experience in the other two new middle schools to validate/extend our elaborated models.

References

[1] Bellatreche L, Garcia F, Pham D N, et al. SONDER: A Data-Driven Methodology for Designing Net-Zero Energy Public Buildings. In To Appear in 22 International Conference on Data Warehousing and Knowledge Discovery, 2020.

[2] Boukhari I, Jean S, Aït-Sadoune I, and Bellatreche L. Therole of user requirements in data repository design. Int. J. Soft w. Tools Technol.Transf.20, 1, 2018, 19–34.

[3] Charlet J. Knowledge Engineering: developments, results and perspectives for Knowledge Management in medical field. Pierre and Marie Curie University, 2002.

[4] Reimann R, Cronin D, Noessel C, et al. About Face: The Essentials of Interaction Design, 2014 (4th. ed.).

[5] Dabbebi I, Iksal S, Gilliot J.M et al. Towards Adaptive Dashboards for Learning Analytic - An Approach for Conceptual Design and Implementation. In Proceedings of the 9th International Conference on Computer Supported Education, 2017, 120–131.

[6] Vondra P, Divjak B. Learning Analytics: Meeting the Needs of Students and Teachers in Pre-tertiary Education. In Central European Conference on Information and Intelligent Systems, 2017, 117–250.

[7] Bellotti V, Dourish, P. Awareness and Coordination in Shared Workspaces.1992, 107–114.

[8] Engeström Y. Learning by Expanding An Activity-Theoretical Approach to Developmental Research. Cambridge University Press, 1987.

[9] Gilliot J.M, Iksal S, Medou D.M et al. Participatory design of learning analytics dashboards. In IHM, 2018, 119–127.

[10] Lebis A, Lefèvre M, Luengo V et al. Capitalisation of analysis processes: enabling reproducibility, openness and adaptability thanks to narration. In8th International Conference on Learning Analytics and Knowledge (LAK), 2018, 245–254.

[11] Long P. D, Siemens G. Penetrating the Fog: Analytics in Learning and Education. Educause Review. EDUCAUSE review46, 5, 2011, 31–40.

[12] V. Luengo. What indicators to personalize learning in a digital school? 2016. https://cache.media.eduscol.education.fr/file/Numerique/48/2/rapport_QuelsIndicateurs_LIP6_612482.pdf

[13] Mayring P. 2010. Qualitative In halts analyse. In Handbuch qualitative Forschung in der Psychologie, 2010, 601–613.

[14] The Ministry of National Education (France), 2020, http://www.education.gouv.fr

[15] Pierrot L, Cerisier J-F, El-Kechaï H et al. Using a Mixed Analysis Process to Identify the Students' Digital Practices. In 12th European Conference on Technology Enhanced Learning EC-TEL, 2017, 448–453.

[16] Reffay C, Betbeder M.L, Chanier T. Multimodal learning and teaching corpora exchange: lessons learned in five years by the Mulce project. International Journal of Technology Enhanced Learning, 2012,11–30.

[17] Schwendimann ,B.A, Rodríguez-Triana M.J et al. Per-ceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research. IEEE Trans. Learn. Technol.10, 1, 2017, 30–41.

[18] Vinck D. Taking intermediary objects and equipping work into account when studying engineering practices. Engineering Studies 3, 1, 2011.

[19] Wilms K.L, Meske, Stefan Stieglitz C et al. Digital Transformation in Higher Education - New Cohorts, New Requirements?. In 23rd Americas Conference on Information Systems, AMCIS, 2017.

First A. Author Photo. Biographies should be limited to one paragraph consisting of the following: sequentially ordered list of degrees, including years achieved; sequentially ordered places of employ concluding with current employment; associa-tion with any official journals or conferences; major profes-sional and/or academic achievements, i.e., best paper awards, research grants, etc.; any publication information (number of papers and titles of books published); current research interests; association with any professional associati