

# Learning Analytics Deployment at an University

Elisângela Mônaco de Moraes<sup>(✉)</sup> and Márcia Terra da Silva

Paulista University, São Paulo, Brazil  
emonaco@unip.br

**Abstract.** This article presents the implementation of Learning Analytics in a University with the aim of testing the extraction of knowledge from databases of two systems used in distance education, aiming to provide support for the management of the distance education as a way to understand students' difficulties. Therefore, it was validated the applicability of literature reference model for the implementation the Learning Analytics. The data were extracted from the system of academic management and from the entrance exam to universities and were chosen the integration data, database and analytical tools. On the presented results, the model has been successfully applied, guiding the implementation of Learning Analytics at the University. The chosen tools facilitated the deployment and, brought benefits to University.

**Keywords:** Learning Analytics · Tools visualization · Distance Education

## 1 Introduction

The Distance Education is a type of education in expansion, which meets the new demands of a society oriented to information, as in new lifestyles and of consumption. We live in an interconnected world over the network, where more and more people study at home, as they can access from there, the information available online.

The distance education in Brazil is seen as a potential model to enable the demand for higher education in Brazil. EAD democratizes the entry to higher education, for the portion of the population that did not have undergraduate course offering in their localities. The expansion of this modality is clearly in higher education census conducted by INEP.

According to the technical summary of 2014 higher education census, published by the National Institute of Educational Studies Anísio Teixeira [1], which provides information on higher education in Brazil, the distance mode continues growing, with 1.34 million enrollment of higher education, representing 17.1% of total enrollment. Noteworthy the degree, where there were a total of 1,466,635 enrollments, with 540,693, or 36.9% of the distance mode. This figure shows an increase of 6.7% between 2013 and 2014.

This increase in the number of students in the distance education mode, and the characteristic of this type of education which is measured by Information and Communication Technology, results in an increasing volume of data produced,

© IFIP International Federation for Information Processing 2016

Published by Springer International Publishing AG 2016. All Rights Reserved

I. Nääs et al. (Eds.): APMS 2016, IFIP AICT 488, pp. 79–85, 2016.

DOI: 10.1007/978-3-319-51133-7\_10

introducing the need to manipulation of various types of data in a fast way. Thus universities, seek ways to manage these data in order to transform them into information, and with that information, take faster decisions so they can be effective in their actions in the various organizational levels.

In this sense the objective of this study is to verify whether it is possible to extract knowledge from the two systems databases that are used in distance education, the main system that is the Academic Management and the Vestibular system. Therefore it will be validated the applicability of the reference model of Chatti et al. [2] chosen in the literature review of Moraes et al. [3] so that it is possible the extraction of knowledge from these two databases. Thus, provide support for the management of the Distance Education, in order to improve the targeting strategies for the achievement of goals, for example the evasion control.

## 2 Learning Analytics

Different definitions are assigned to the term learning analytics, being one of the most widely adopted in the literature: “the measurement, collection, analysis and reporting of data of students and its contexts, for the purpose of understanding and learning, optimizing the environment it occurs in” [4]. Another definition for Learning Analytics is that if applied in several organizational levels, where each level gives access to a different set of data and contexts, this process provides valuable information [5].

## 3 Methodology

This study used data obtained from the Information Systems used in Distance Education, specifically the Academic Management System and Vestibular system. The courses that will be analyzed in this research are technological undergraduate courses, both information and communication area, as follows: Course of Technology Analysis and Systems Development, whose mnemonic is DS, this course is offered in the distance model at the University since 2014 and now has 1,889 students; and the Course of Technology in Information Technology Management, whose mnemonic is IT, this course is offered in the distance model at the University since 2009 and today has 2,122 students. So that it is possible the realization of this work, the research was divided into distinct phases: research questions; collection and production of data; search results and conclusion.

### 3.1 Application Process of Learning Analytics

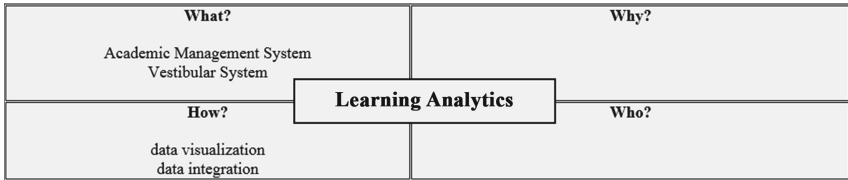
The reference model, proposed by Chatti et al. [2], divides the application of learning analytics in 04 dimensions: what? Why? Who? How?, thus facilitating and classifying the literature. Hence, using the model for the application of Learning Analytics, the following research questions will be addressed:

- RQ1 - How and what are the data to be analyzed?
- RQ2 - What are the aims of the analysis carried out and who will be presented?

## 4 Collection and Production Data

RQ1 - How and what are the data to be analyzed? Using the reference model Chatti et al. [2] were chosen tools and which systems used for this analysis. Thus, to achieve the goal of dimension How?, proposed in the model, the tools will be used the tools:

- Talend Open Studio - An open source solution for data integration and extraction, transform and load (ETL)
- Database HP Vertica - Database to streamline the handling of high volumes of data, its use is free up to 1TB of data, which, for this research is well above the need of analysis.
- Tableau - Data analysis tool, leader in the market, has a public version and is also free for teachers, easy to use. The choice of systems for the collection and analysis of data, thus meeting the dimension What? (Fig. 1), was carried out by the researcher, based on the analysis aim, namely:
- Academic Management System - Responsible for controlling the administrative processes and the academic management of the University
- Vestibular System - system responsible for the student, admission to the University.



**Fig. 1.** Learning analytics (Source: Adapted from Chatti, and Dyckoff Thus)

The tools were chosen by the large volume of data, around 100 million records only in the Academic Management System, the variety of data and the speed required in the presentation of information. Each system uses a transactional database, but using a columnar database connected to a display tool, we obtain some benefits as: reduced time for consultations; ability to analyze faster the University data and thus consolidating the information into a single tool for use in the various levels of the organization (Fig. 2).

RQ2 - What are the aims of the analysis carried out and who will be presented? Using the reference model Chatti et al. [2], dimensions, Why? and Who? were filled with questions asked by the course coordinators, namely: How are the students distributed supported poles and what is the profile? What are the subjects that failed most? What is the number of students who pass the entrance exam? The managers of the University would like to know way online: How many people enrolled in the University entrance exam, by year, month, and day? What is the course that has the largest number of students enrolled? What is the profile of the student who seeks the IES, to take the entrance exam? This



Fig. 2. Consolidating information

<b>What?</b> Academic Management System Vestibular System	<b>Why?</b> How are the students distributed supported poles and what is the profile? What are the subjects that failed most? What is the number of students who pass the entrance exam? How many people enrolled in the University entrance exam, by year, month, and day? What is the course that has the largest number of students enrolled? What is the profile of the student who seeks the IES, to take the entrance exam?
<b>How?</b> data visualization data integration	<b>Who?</b> Coordinators Course managers

Fig. 3. Learning analytics (Source: Adapted from Chatti, and Dyckoff Thus)

way, the model was completed in all dimensions, as Fig. 3, completing with the following information the dimensions, Why? and Who?

Now with the model completed in its four dimensions, the results will be presented to the Coordinators Course and managers of IES.

4.1 Results

In this section, it will be presented the results of analyzes, based on the questions raised, first by the coordinators and then by managers. For the data to be presented on online form to those involved, data integration tool was connected to transactional databases, data were extracted, transformed and loaded into columnar database and visualization tool connected on this data-base, so that it were developed the visualizations. Information is updated hourly and the involved people received access credentials.

For the coordinators the information's were presented, first answering the question: How are the students distributed at the support poles and what is the profile? (Fig. 4).

The information extracted for the coordinators of the two courses, show the highest concentration of students in the Southeast, the number of students per pole, the percentage of students by sex, being the male far superior, to DS course 88.06% are male and 11.94 are female, and for IT course are 13.62% female and 86.38% male. The number of enrolled students, distributed by age group for both courses, has the highest concentration in the range of 30 to 34 years.

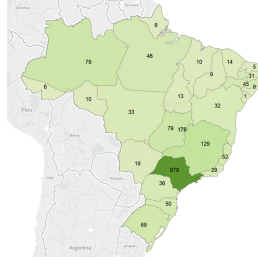


Fig. 4. Geographical distribution of students

In the visualization, the coordinators go through the information, interacting in a drill-down fashion, thus, down in the details as follows: in the student enrollment numbers information, you can click to go in the geographical distribution by state, clicking again appear the information by city and in the city, the presential support poles, in the poles appear the students and clicking in the student is possible to view the failed, approved, registered, exempt and locked disciplines.

For questions: What are the disciplines that most failed? What is the number of students who pass the entrance exam? The visualizations were presented as Figs. 5 and 6.

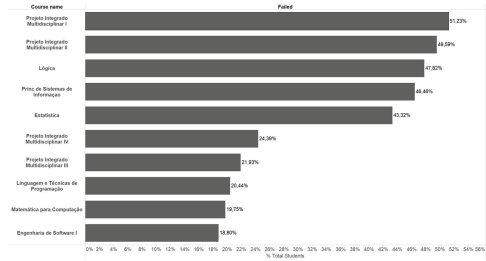


Fig. 5. Failed in the disciplines and exam

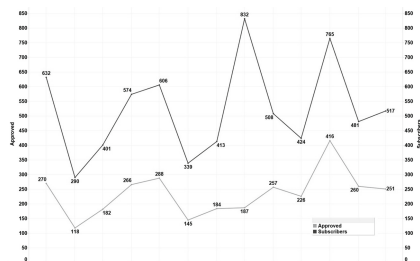


Fig. 6. Number of students who pass the entrance exam



- Reduced time to check the number of entries in the vestibular system, because the data were processed at dawn and presented in spreadsheets, which with the deployment the information's are accessed online.
- The access to information from different systems, concentrated in a single tool.

By gaining the ability to analyze students from the extraction of knowledge, using the integration, database and visualization tools, the University gained a clearer profile of their student population. Managers and Coordinators of courses involved in this study evaluated that the information's presented are useful and relevant. Since the focus of this study was the implementation, from a model of learning analytics, and thus extracting knowledge from data, in the next studies are suggested the correlation of data from other systems. Additionally it is necessary expand to more departments of the higher education institution.

## References

1. Instituto Nacional de Estudos e Pesquisas Anísio Teixeira. <http://portal.inep.gov.br/basica-levantamentos-acessar>
2. Chatti, M.A., Dyckhoff, A.L., Schroeder, U., Thüs, H.: A reference model for learning analytics. *Int. J. Technol. Enhanced Learn.* **4**(5–6), 318–331 (2012)
3. Moraes, E., Silva, M., Souza, M.: Models to implement learning analytics: a literature review. In: 27th POMS Annual Conference, Orlando (2016)
4. Siemens, G., Long, P.: Penetrating the fog: analytics in learning and education. *Educase Rev.* **46**(5), 30 (2011)
5. Shum, S.B., Crick, R.D.: Learning dispositions and transferable competencies: pedagogy, modelling and learning analytics. In: 2nd International Conference on Learning Analytics & Knowledge, British Columbia, Vancouver (2012)