Students’ perceptions of academic performance in Distance Education evaluated by Learning Analytics and Ontologies

Laécio A. Costa¹, Marlo Souza², Laís N. Salvador², Aleph C. Silveira³, Celso A. Saibel³

¹Informatics Collegiate - Instituto Federal do Sertão Pernambucano (IFSertão) Petrolina, PE – Brazil *
²Computer Science Graduate Program – Federal University of Bahia (UFBA) Salvador, BA – Brazil
³Department of Informatics – Federal University of Espírito Santo (UFES) Vitória, ES – Brazil

laecio.costa@ifsertao-pe.edu.br, {msouza1, laisns}@ufba.br, 
{aleph.campos, celsoalbertosaibelsantos}@gmail.com

Abstract. Keeping track of student learning progress in Distance Education courses is a challenge for education specialists. It is not always possible to adopt the same evaluation mechanisms employed in traditional education due to transactional distance between educators and students. To help educators in the process of evaluating the acquisition of skills and monitoring students’ learning experiences, we present the SapeS architecture, which employs Learning Analytics and Ontological Reasoning to produce information about student performance. This architecture enables both educators and students to intervene in the teaching process and promotes autonomy to students successfully achieve their planned objectives. This proposal was based on the Design Science Research methodology, which enables the production and evaluation of the proposed solution for better understanding of its viability in the participant’s real context.

1. Introduction

The use of Information and Communications Technology (ICT) in education has been growing worldwide due to the confluence of new technologies, the global adoption of the Internet, and the demand for a workforce trained for the constantly evolving digital economy [Palvia et al. 2018]. These technologies enable advantages that include breaking barriers to the asynchronous and synchronous learning network model, accessing teaching materials and reducing costs with classrooms and equipment. This modality in which educational content is delivered and administered over the Internet is called Distance Education. Distance education has brought new educational challenges such as the lack of reliability, transparency, and feasibility in assessing student performance since many traditional assessment methods are not suitable in this modality [Yago et al. 2019]. Therefore, the use of Learning Management Systems (LMS) has become widespread and a primary concern as educational institutions seek to disseminate knowledge and educational content with Distance Education regardless of the challenges mentioned above.

*Acknowledgments to the Federal Institute of Sertão Pernambucano (IFSertão-PE).
As teaching and educational technologies tend to be adaptive, due to the transactional distance between students and educators in Distance Education, new methods of assessing academic performance are needed for these environments. However, for [Cabrera-Loayza et al. 2020], methods applied to monitor and evaluate students’ academic performance in traditional classrooms are neither effective nor appropriate for application in Distance Education.

[Fotiadou et al. 2017] argue that educators must guide students according to a pedagogical plan that allows each student to build an autonomous learning process based on clear educational objectives to promote gains in learning. According to [Bloom et al. 1956], and [Freire 2002], Educational Objectives or Learning Objectives (LO) are resources for pedagogical use, inherent to the teaching process, that guide the planning of the course and allows evaluation of the student’s academic performance. Student’s can monitor the fulfilment of the planned educational objectives, enabling self-evaluation and, consequently, the self-regulation of their learning process.

Students’ knowledge acquisition occurs through the learning experiences carried out by interactions with the pedagogical activities (or Learning Units, LU) made available in the LMS. The educator, therefore, plans the LU to build knowledge or skills in the student. Observing the learning experiences, the instructional sequencing, and the students’ learning path allows them to monitor the students’ planned skills development and achievement of learning goals. The learning path is planned based on the skills and competencies that the student is intended to acquire in their classes, each of them associated with one or more LU.

In Distance Education, educators try to use or adapt the LMS’ tools to monitor and evaluate their students’ academic progress. However, some authors [Palvia et al. 2018] argue that monitoring the students’ academic performance in this context is difficult, as it is hard to obtain the necessary information for such in the LMS. According to [Fotiadou et al. 2017], the LMS, while still a rich source of educational data, does not provide tools for an efficient analysis of student progress since this plethora of information in the LMS is of great value to teachers and can be used to analyze students’ academic progression, predict behavior and measure their performance [Avella et al. 2016].

Studies about Distance Education and LMS were identified by a systematic literature review conducted by [Costa et al. 2020]. None of the studies found by the authors addressed the integration of Learning Analytics (LA) methods with information about the pedagogical structure of an online course, formalized through Taxonomies of Educational Objectives (TEO). This integration allows a more nuanced evaluation of students’ academic performance, as it is connected with the learning objectives for which the course is designed. Nevertheless, LA technologies have been employed before in the literature to help the evaluation process in Distance Education, and some approaches should be highlighted.

Firstly, [Yago et al. 2019] proposed a model to identify the student’s profile based on a network of ontologies to offer educators several indicators about the student’s profile. This proposal, called ON-Smile, aims to provide educators with the student’s academic status to improve their supervision. [Cabrera-Loayza et al. 2020], on the other hand, describe a tool to present students’ progress, highlighting those at risk of abandon-
ing the course. This LA-based tool allows dynamic visualization of educational information from interaction events discovered in the learning environment log files. Lastly, [Nussbaumer et al. 2012] present a conceptual architecture to detect and analyze cognitive learning activities in personal learning environments. The architecture aims to identify non-observable cognitive behavior and provide feedback to the students to raise awareness of their own cognitive processes that cannot be observed directly in the environment.

It is important to note that these previous proposals differ from our work in numerous ways. ON-Smile [Yago et al. 2019], for example, does not relate students’ state of knowledge to the curriculum and the LO planned by educators. The LA-based tool proposed by [Cabrera-Loayza et al. 2020] does not promote the triangulation between the student’s academic situation with the learning experiences carried out and the fulfillment of educational objectives to assess the educational situation, as our proposal does. Finally, [Nussbaumer et al. 2012] conceptual architecture does not propose verifying the acquisition of knowledge related to the educator’s planning, although the authors state that the results can be applied in conjunction with LA techniques. In short, these studies lack i) parameterization through ontologies that allow the educator to specify the taxonomic structure of educational objectives in the course, ii) feedback to educators and students, and iii) evaluation of the academic performance through the inference of information guided by LO Taxonomic structure.

Therefore, this paper presents the SapeS architecture (Student Academic Performance Evaluation System architecture) based on the use of LA and the formalization of TEO by means of a computational ontology to aid the evaluation of students’ academic performance in Distance Education. According to [Gruber 2009], an ontology specifies a vocabulary of assertions, which can be inputs or outputs of knowledge agents, such as a software program. As an interface specification, an ontology provides a language for communicating with the agent. An agent that supports this interface does not need to use the ontology terms as an internal encoding of its knowledge. In this work, SapeS tool aims to extract information about academic progress and provides evaluations of students’ academic performance based on the LO planned by the educator. Data on learning experiences are collected from the LMS and analyzed based on the formalization of the TEO in an online learning environment so academic performance can be measured. A set of information is then available to the educator and the students to assist in measuring the acquisition of skills and competencies.

The research question of this work is: Q1: Is SapeS architecture capable of promoting student’s academic performance based on Learning Analytics and Ontologies, and Q2: SapeS presenting adequate information that promotes gains in teaching and learning process based on Learning Units? This work presents the proposed architecture and the method to verify its viability in a real environment, with volunteer students and teachers from a master’s course in Professional and Technological Education from a Brazilian Federal Institute of Technological Education.

The sections are organized as follows: In Section 2, we detail the methodology of our study and the resulting artifact; in Section 3, we present an empirical study of the applicability of the proposed architecture by a user evaluation, discussing the methods and materials employed in the study; in Section 4 we present the results of the empirical study.
as well our discussion. Finally, Section 5 presents the conclusions and future works.

2. Design Science Research

In this section we present the methodology Design Science Research (DSR) and our proposed artifact. DSR is defined as a set of guidelines and methods for the process of creating, building, and validating observed problem-solving artifacts, contributing to research, evaluating projects, and communicating the results to the appropriate audience [Wang and Wang 2013]. The term Design Science Research first appeared in the 60s with [Fuller 1965] and [Gregory 1966], as both authors agreed on the need to look for a more systematic way of designing artifacts. An artifact is designed for a purpose, such as an everyday object like a pencil, a house, or a car, and they are not restricted to physical objects but can also be intangible intellectual devices, such as an educational activity or methodology. During the development of this artifact, new problems and knowledge are discovered from its practical application. Thus, considering a cyclical method of development and application, this artifact is presented in the next subsection.

This methodological approach was adopted in this research to investigate the viability of the SapeS since it allows the creation and evaluation of artifact(s) to solve identified problems. With DSR, it is also possible to trace a path that provides understanding and approach to the problem raised. The methodology is structured in five steps of actions, as shown in Figure 1.

Based on this structure, the currently step of this work is situated between Step 4. Evaluation and Step 5. Conclusion.

2.1. The artifact: Sapes Architecture

This architecture was developed to collect learning experiences in the LMS and evaluate academic performance based on the planned LOs, employing information obtained through reasoning based on the OntoLO ontology [Costa et al. 2018], and [Costa et al. 2019]. SapeS was developed from the reference model for LA applications proposed by [Cho and Lee 2016] and ISO/IEC-TR 20748-1 [TR20748-1 2016], both identified in the systematic review aforementioned. Figure 2 presents the expanded ISO/IEC-TR 20748-1 with Verification Data Compliance and conversion to RDF as well as data processing with OntoLO.

---

1From the Latin *arte* and *factus*, means made with art, with technique. It can be defined as an artifice, something constructed artificially, intentionally, has a specific utility, it is the result of a project.
The SapeS architecture, presented in Figure 3, is composed of three modules, consisting of the Educational data collection and store, which interfaces the LMS and the reasoning services, the Processing and analysis service, which performs the reasoning and data processing steps, and the Dashboard, which presents the interface of the system to the users.

Module 1 (Data Collection) is responsible for collecting, verifying compliance, adding semantics, and storing data by making it available in a Learning Record Store (LRS). The architecture was designed so that it is possible to use different Learning Activity Sensors (LAS) to collect data from different learning environments. In this work, for data collection, it was adopted the xAPI framework plugin for the Moodle environment.

Applying sensors for data collection implies the need for a step for data compliance verification. This stage is responsible for receiving the declarations in triple format (Actor-Verb-Object), checking if they conform to the xAPI specification, transforming the data to the RDF/XML format, and storing them in the repository (Learning Record Store). Since the collection can generate a large volume of data, the elements collected are stored in a NOSQL database, following the concepts of Data Lake [Inmon 2016]. This ensures that all student data sent by the sensors will be stored according to the xAPI specification, avoiding losses that could impact the architecture’s operation.

Module 2 (Processing and Analysis) is responsible for consulting the LRS with data on students’ learning experiences and sending them to the Inferences stage. In the Inferences stage, the SapeS architecture is parameterized through an ontology that formalizes a taxonomy of LO. This ontology is populated with the activities (LU), expected

---

2Learning Locker was adopted as a repository for being open source and distributed under a GPL3.0 license (https://www.ht2labs.com/learning-locker-community/overview/).
3https://adlnet.gov/projects/xapi/
4Environment used in the teaching institute in which the investigation was carried out.
5NOSQL is a generic term that represents non-relational databases.
6Data Lake is a repository of data stored in its natural/raw format.
Figure 3. Overview of the SapeS architecture.

skills, and respective learning objectives planned by the educator. It produces inferences from the learning experiences collected during the learners’ interactions in the course and stores them in the SapeS database. The inferences and information extraction will provide additional information for analysis through statistical algorithms presented in the next stage. For this work, the OntoLO ontology was used as a basis for generating knowledge concerning the cognitive domain of Bloom’s Revised LO taxonomy. OntoLO is currently public through GitHub\(^7\) and can be view and edited with Protégé\(^8\) and WebVOWL\(^9\) tools.

The information analysis stage is responsible for classifying and measuring each student’s academic performance, according to the learning experiences collected and the inferences made. With the educational data stored and available, two statistical algorithms, Linear and Logistic Regression algorithms, were used to analyze a dependent variable (academic performance) and its relationship/association with another independent variables and training values (learning experiences, inferences, and scores). The linear regression algorithm aims to predict a numerical value (academic performance) from educational data. The logistic regression algorithm seeks to predict values based on a categorical binary variable (for example, if the student’s situation is favorable or not in his academic goal) according to a series of explanatory variables. This analysis is referred

\(^7\)https://github.com/laeciocosta/ontolo

\(^8\)Protégé is a free, open-source ontology editor and a knowledge management system developed by Stanford University

\(^9\)Web application for the interactive visualization of ontologies.
to as an instrument that performs the analysis of academic performance using indicators of academic progress in Distance Education [Yamada et al. 2016][Li et al. 2017].

Module 3 (Presentation of information) is responsible for providing graphic reports (educational dashboards) for the educator and the student’s profile. As LA is a cyclical process, the dashboards provided in Module 3 enable the educator to abstract information, carry out analyses, and change pedagogical strategies to promote gains in learning and then feedback the LA process. Students can monitor their performance and follow their evolution according to the educator’s plan, promoting their engagement and self-regulation learning.


To evaluate and understand this architecture’s applicability in an educational context, an investigation was conducted in a course of a Master’s program in Professional and Technological Education class from May to July 2020. The purpose of this investigation is to verify the feasibility of the proposed architecture, noting its impacts on the teaching and learning process.

Initially, the environment was prepared with the LRS\textsuperscript{10}, the SapeS Web Tool\textsuperscript{11}, and a source of semantic data available in the LRS. The sensors for data collection were installed and configured in the LMS and the educators were trained in the use and functionality of SapeS. The academic course observed used the LMS Moodle, and all students’ learning experiences occurred in this environment.

The LRS received more than 10,000 declaratives, i.e. interactions of educators and students with the LMS, through the execution of the course. As the educator adds new learning units and students’ interactions with the LMS occur, the LAS sensor captures and sends the declarations to the Data Collection Module. In order to avoid overloading our data traffic.

During the investigation, the evaluation of the SapeS-based tool was carried out by the participants. Initially, interviews with educators were carried out to obtain information on their perceptions about the evaluations strategies and academic performance of their students, in addition to obtaining information about the use of pedagogical tools and resources to monitor the expected knowledge acquisition by the students. We also conducted interviews to collect qualitative information (cognitive and behavioral) linked to positive and negative perceptions about the use of information produced by the SapeS tool. We sought to identify the degree of acceptance of participants regarding the characteristics of the proposed tool and its application in the process of teaching. In addition, we obtain information about the perception of the tool’s ease of use, the assistance provided to the process of evaluating students’ academic performance, and its usefulness in control actions to assist in the students’ learning process.

Two questionnaires were applied through Google Forms to diagnose and understand the student’s understanding regarding the object of this work. After the student volunteers experimented with the SapeS tool, another questionnaire was applied to identify the volunteers’ experiences regarding the SapeS functionalities and usability, the value

\textsuperscript{10}http://lrs.sapes.cc/
\textsuperscript{11}http://www.sapes.cc
of the information displayed on their dashboards, and its viability. In short, in the first questionnaire, the volunteer students reported about their experiences and lack of Moodle tools that allow monitoring of academic progress and encourage student involvement, and by the end of the course, an usability questionnaire about the information provided by the SapeS web tool was applied.

The first questionnaire concerns student engagement and self-regulation of learning in Distance Education and contains 61 questions. The second questionnaire was related to the usability of the SapeS tool and contained 11 questions. Both questionnaires follow the Likert scale. Cronbach’s alpha's coefficient was applied to verify the reliability, consistency, and correlation of the answers to the applied questionnaires. The closer Cronbach’s alpha reaches 1, the greater the reliability of the questionnaire. The questionnaires applied in this work achieved alpha values of 0.965 and 0.944, respectively.

At the end of the course, interviews were conducted to collect data to carry out a qualitative analysis to measure the impacts that the tool had in the promotion of learning. The objective of the interviews was to analyze the data, from the perspective of students participants, about the factors that determine their engagement and how academic performance interferes with self-regulation of learning.

4. Results and Discussion

In general, results revealed that the interactions in Moodle, the cognitive level inferred by the OntoLO ontology, the concept obtained in each evaluative pedagogical activity (LU), and the interaction in the LMS had significant effects on the student’s academic performance. In addition, the possibility of adapting the didactic-pedagogical actions changed the student’s behavior being identified through of the interactions grow with the LU’s in Moodle and their interactions with the SapeS tool. Interestingly, the explicit disclosure of the educational objectives and the expected cognitive ability related to each activity demonstrated the importance of these elements in the learning process, which contributed to a better understanding of the course’s planning. Many students reported their use of the SapeS web tool to monitor their activities and academic performance compared to other colleagues, according to Figure 4. This process favored searching and sharing information among colleagues to reach the expected cognitive level for the planned activities.

As for the data obtained from the questionnaires, the average score was calculated according to the Likert psychometric scale, with the following conditions being attributed to the average scores: values less than 3 represent the perceptions of disagreement (unfavorable), values equal to 3 “indifferent” (neither agree nor disagree), and values above 3 represent perceptions of agreement (favorable). The results of the Engagement dimension achieved an average of 4.34, indicating a good perception about students’ engagement categorized through pre-conditions in the academic background, aptitude, individual student characteristics, and a structure of the course that allows the perception of an adequate level of engagement. As for the dimension of Students’ academic progress, this had an

---

12 The questionnaires can be accessed via: https://cutt.ly/NkTHheo
13 The Likert is a type of psychometric response scale in which respondents specify their level of agreement with a statement, usually in five points: (1) Strongly Disagree-SD; (2) Partially Disagree-PD; (3) Neutral-N; (4) Partially Agree-PA; (5) Strongly Agree-SA.
The interviews carried out allowed us to identify elements such as: (1) Ease of use; (2) Quality of the information presented; (3) Difficulties encountered; (4) Impacts provided on the learning process (engagement, motivation, and self-regulation); (5) Suggestion of features to include in the proposed tool.

The data collected revealed an above-average positive perception about the usability of SapeS tool and its impacts on learning. In this dimension, most students (70%) had a positive perception related to satisfaction, ease of use, and impacts on learning provided by the SapeS tool.

Students were unanimous in reporting that the information provided by SapeS tool is clear, objective, and coherent with the planned pedagogical actions by the educator. Also, with the SapeS-based tool, monitoring their educational path based on competencies was effective as they realized which skills they needed to devote more time and attention to achieve the Educational Objectives planned by the educator.

Some benefits identified according to the participants’ assessments are:

1. Relevance of the architecture and agreement with the adopted procedures: 100% of the professors agreed with the importance of the proposal as an aid to the educator in the process of evaluating academic performance and to the student as a way of supervising the academic path in accordance with the educational objectives;
2. Tool usefulness: 100% of the participants consider the solution useful for the process of evaluating the students’ learning path;
3. General aspects of the tool: The tool was classified by the participants as adequate
Another factor observed is that it is essential to analyze the educator’s feedback (e.g., grades) of the students’ activities, as at times the educators introduced a new LU without delivering feedback on activities from previous units on which it depends - in this case, the dependency between the LU is an important factor, as the course planning can be non-linear. According to [Yamada et al. 2016], the learning strategy planned by the educator based on educational objectives contributes to the improvement of student behavior in the pursuit of knowledge and improvement in learning performance. Thus, with the delay in delivering feedback from previous activities, students were unable to assess their real performance and seek ways to improve skills to progress in the course, impacting their overall performance throughout their educational path.

5. Conclusion and Future Works

This article presented the SapeS architecture, which allows the monitoring and evaluation of students’ academic performance in Distance Education based on LO Taxonomies. This architecture correlates students’ learning experiences with LO, and LU and to accomplish this, it associates ontologies with Learning Analytics. The ontological formalization of the Taxonomy of Educational Objectives allows inferring pedagogical information about the planning of didactic-pedagogical actions and also enables the software agent to understand abstract concepts of the referred taxonomy, relating them to the planned Learning Unit. This interdependent collaboration between LA, LU and LO is the main argument of this research and reveals a wide field of study.

This architecture offers various contributions in the educational context, such as monitoring students’ academic performance, assisting in the process of self-regulation of learning, and, consequently, promoting gains in the teaching and learning process. It allows educators to assist in the process of assessment and pedagogical adequacy of instructional sequencing. Educators can better understand and follow their students’ learning, taking into account their deficiencies and needs.

Based on the results of the investigation, the amount of educational data collected is a rich source of knowledge that can help improve all aspects of Distance Education. The available computational techniques and processes can assist in exploring this data, and when allied with pedagogical instruments for measuring the acquisition of knowledge, it is possible to carry out an assessment of the student’s academic performance in line with the planned didactic-pedagogical actions by educators. Some limitations were observed due to the partial results obtained, such as the educator’s time to complete the assessments and make the analysis/feedback of evaluative activities available. As grades are performance indicators, delays have consequences for the analysis in real-time due to learning sequencing.

Future works intend to mitigate these limitations by notifying the educator about activities pending assessment and recommending the formation of heterogeneous groups of students in the class to balance learners’ performance. Such actions can improve the assessment process and further motivate students to self-regulate learning.
References


