ABSTRACT

This paper describes the architecture of a web application developed and deployed based on Learning Analytics techniques, with the aim of supporting the academic decision making required of educational administrators, on behalf of students. Our work in this field involves a warehouse approach to managing the information supported by the platform. The data integration is modular, allowing the introduction of new databases through the use of a middleware platform, WSO2 Data Services Server, which takes data from different silos and makes them available as a set of WS style web services. The three layer web architecture model is explained in detail, and the subsequent deployment of the web application called UDLearn is shown along with the primary results obtained from its implementation at one of the largest public universities in Colombia, which are shown as a toolkit to help administrators to improve the impact of their decisions in the academy community.

KEYWORDS

Learning Analytics, Software Architecture, Data Mining (DM), Human Computer Interaction (HCI), WSO2 Data Services Server.

INTRODUCTION

One of the major challenges that software engineers face, is to generate useful products that provide effective solutions to problems presented in a specific context. Thus new techniques are always tested and included in new architectures to enhance their functionality and diminish the efforts required by the user.

Learning Analytics is an emerging field in which sophisticated analytic tools are used to improve learning and education. It draws from, and is closely tied to, a series of other fields of study including business intelligence, web analytics, academic analytics, educational data mining and action analytics [1].

Several research papers have used Learning Analytics (LA) in a distance education context, having in mind that LA analyze interaction between the students with online educational resources, which make it primarily of use to institutions that offer distance education. Aside from this, the data processed by this tool comes mostly from Intelligent Tutoring Systems (ITS) and Learning Management Systems (LMS), without correlating additional information gathered in other systems that could also be relevant in order to get a fuller picture of the student learning experience.

UDLearn is developed and deployed with the aim of supporting the decision making process of administrators and faculty decision makers, analyzing academic activities in a conventional educational context (face-to-face learning education model) through a web application that integrates data from LMS, administrative institution databases and other administrative entities inside the university, thereby offering the necessary functionality to make decisions on the behalf of the students.
1. LEARNING ANALYTICS

A reference model for LA by Dyckoff et al., [2] claimed that: “The Learning Analytics process is often an integrative cycle and it is generally carried out in three major steps: Data collection and preprocessing, Analytics and action, and post-processing.” The reference model is based on four dimensions as shown in *figure 1*.

1.1 Data Environment (What?)

LA is a data driven approach. LA approaches use varied sources of educational data. These sources fall into two big categories: centralized educational systems and distributed learning environments. Centralized educational systems are well represented by learning management systems (LMS).

1.2 Stakeholder (Who?)

The application of LA can be oriented toward different stakeholders, including students, teachers, (intelligent) tutors/mentors, educational institutions (administrators and faculty decision makers) [3], researchers, and system designers with different perspectives, goals, and expectations of the LA exercise.

1.3 Objectives (Why?)

There are many objectives in LA according to the particular point of view of the different stakeholders. Possible objectives of LA include monitoring, analysis, prediction, intervention, tutoring/mentoring, assessment, feedback, adaptation, personalization, recommendation, and reflection.

1.4 Methods (How?)

LA applies different techniques to detect interesting patterns hidden in educational data sets. In this section, we describe four techniques that have received particular attention in the LA literature in the last couple of years, namely statistics, information visualization (IV), data mining (DM), and social network analysis (SNA)”.

2. SURVEY

Learning Analytics have been used to understand and analyze the learning process based on the static and dynamic data extracted from virtual learning environments [4]. The incidence of studies in the area has increased rapidly since 2010, they target intelligent tutors in adaptive systems (48%), researchers/system designers (30%) teachers and students (17%) and educational institutions (5%) [2].

Interesting examples come from prestigious Universities such as Purdue University [5], Baylor University, University of Alabama, Northern Arizona University, Sinclair Community College, among others, which based on LA studies in their communities have changed decision making, planning allocation and student retention, obtaining a remarkable improvement in their students’ degrees [1].
Projects such as that developed by the Universidad Carlos III de Madrid, show the drafting and implementation of ISCARE[5], a piece of software which uses Learning Analytics to give feedback on learning processes, thereby achieving the following: Identifying the effectiveness of learning elements, increasing student participation in the learning environment, guiding tutors’ preparation of activities, etc. In comparison to this research paper, that approach extracts information from different sources and applies Learning Analytics techniques to analyze it and show it in a web application. The development of ISCARE however is carried out in a virtual educational environment with the aim of generating suggestions and feedback about the leaning process, it does not examine the impact of Learning Analytics in a face-to-face education model.

Aachen University deployed eLAT, a toolkit designed to support and improve the processes involved in online teaching methods based on personal interests and observations [6]. They provided information about their software development such as requirement analysis, user interface design, deployment and implementation. In the latter section they studied information based on four courses that used the Learning Management System (LMS) available at the University, processing in microseconds large data sets with the aim of analyzing a graphic overview executed in real time. The form of the executed eLAT and Learning Analytics techniques are similar to this approach; nevertheless, this toolkit is focused on supporting a different stakeholder (teachers) in a virtual education environment with datasets solely from LMS sources.

Another aspect of the existing research looks at the application of DSS in academic theory and is focused mainly on formulating the general principles and approaches of model-based Decision Support Systems for academic environments. A number of authors have discussed various aspects of DSS and their use within higher education institutions, these included: (1) academic resources, academic advising and course scheduling; (2) resource allocation, budget planning, corporate governance, performance assessment, strategic planning; (3) admission policy, analysis of enrolment demand, capacity management, and enrolment management [6].

It is therefore worth investigating an Educational Decision Support System (EDSS), so called because the application is tightly linked with the characteristics of the environment where the system is deployed and run [7]. Among the existing literature ([8], [9], [10], [11], [12], [13], [14], [15], [16]) there are several examples of EDSS development.

However, none of these implement a Data Warehouse database model or software architecture in their methodology to solve their target problems. In fact, their methodology is centered on dynamic simulation, operation research, statistics and other administrative and industrial engineering tools implemented in the DSS design stage.

Examining research in this field has shown that Learning Analytics have been applied to the analysis in virtual education to monitor intelligent tutors on the web and help students and teachers during the learning process. However there are few research papers written with aim of supporting the decision making in educational institutions. Likewise it is observed that as LA is an emerging field, research about it or its application at Higher Education Institutions has not been published in Colombia. Our approach will incorporate the Learning Analytics techniques in the context of conventional education at the Engineering Faculty of the Francisco Jose de Caldas District University.

3. WEB APPLICATION ARCHITECTURE
The architecture of the proposed web application, called UDLearn, consists of three layer components:

3.1 Information Source Layer

Data sources supply the information needed by the platform however the formats employed may vary significantly [17]. Apart from this, the data is locked away in monolithic application silos. To this end, a mechanism is needed to access relational and other data locked away in databases and make it available. WSO2 Data Services Server provides a platform for integrating data stores and exposing any source of data such as a web service or REST resources [18]. The Engineering Faculty of Francisco Jose de Caldas District University (case of study) has two representative databases, which are the information sources for the web application. Nevertheless, other databases could be included. Figure 3 shows the Information Source Layer.

WSO2 Data Services Server is the open source software used in this project to expose the data contained in the two aforementioned databases, thereafter enabling queries to be made into the databases in order to obtain the information sought. This step allows the deployment of a new database with a warehouse principle structure for the information needed below. Figure 4 depicts an overview of its architecture.

```
Figure 2: Web Application Architecture Model.

3.1 Information Source Layer

Data sources supply the information needed by the platform however the formats employed may vary significantly [17]. Apart from this, the data is locked away in monolithic application silos. To this end, a mechanism is needed to access relational and other data locked away in databases and make it available. WSO2 Data Services Server provides a platform for integrating data stores and exposing any source of data such as a web service or REST resources [18]. The Engineering Faculty of Francisco Jose de Caldas District University (case of study) has two representative databases, which are the information sources for the web application. Nevertheless, other databases could be included. Figure 3 shows the Information Source Layer.

WSO2 Data Services Server is the open source software used in this project to expose the data contained in the two aforementioned databases, thereafter enabling queries to be made into the databases in order to obtain the information sought. This step allows the deployment of a new database with a warehouse principle structure for the information needed below. Figure 4 depicts an overview of its architecture.

```

```

10
```
governance, security, clustering and management capabilities from WSO2 Carbon. Because of the Data Services Server's component-based architecture, developers have the choice to deploy only the required components by adding in and removing features.

Apache Axis2 is the heart of the WSO2 Data Services Server's SOAP processing engine. In the WSO2 Data Services Server, a data service can be summarized into an XML descriptor file written in compliance with Data Services Descriptor Language (DSDL). DSDL is an XML based language defined by WSO2 to write data services. A custom deployer extending the Apache Axis2 deployer framework is responsible for reading this data service descriptor and creating a data service.

The Data Services Server's management console makes comprehensive management, configuration and monitoring possible though a simple, user-friendly GUI. The GUI is done on a layered architecture by separating the backend and frontend concerns. This allows the user to connect to multiple backend using a single GUI console. The Data Services Server also comes with different types of tools for running, testing and monitoring data services”[19].

3.2 Behavioral Layer

Data previously exposed such as a Web Service needs to be unified through a piece of software which transforms the Data Service Descriptor and stores it in a new data repository (Report DB). UDLearn is designed as a web service application, as shown in figure 5, with the aim of: Harvesting, Integration and Storage in a coherent database the reports of the queries previously carried out. Even though, it is not a functional requirement of the software proposed, it is a task executed as a substantial requirement, in order to create a persistent copy of the collected data to fit information into the platform data model, and Process the information that resides in the Report Database to deploy the functions that include:

Analysis and Mining, using Data mining, Business Intelligence and other techniques incorporated to “discover useful patterns or knowledge from data sources”[20], through which users can analyze the available data and search for interesting patterns. The functionality includes a number of Learning Analytics techniques which can be used such as statistics, information visualization (IV), educational data mining (EDM), and social network analysis (SNA) among others.

Reports Generating, which allows users to find data of interest by way of a search mechanism to navigate through the target information using the indicators constructed for this purpose.

Visualization, allowing, through a friendly visual form, the representation of data to comprehend and make use of hidden behavioral patterns, facilitate interpretation of the educational data, expose problems, promote hypothesis formation and allow the emergence of unanticipated properties, among others.

![Figure 5: Behavioral Layer](image)

3.3. Client Layer

At the Client Layer UDLearn Platform is deployed in the form of a portal offering end users (educational administrators and decision makers) the ability to sign in and subsequently access all services available in the behavioral layer. Figure 6 shows the Client Layer.
The client layer is based on Human Computer Interaction Techniques, “which are important in visualization, because a good combination of them can help users to design a good visualization system while optimizing its visualization effects”[21]. Hence, UDLearn achieves the principles of an interactive product design to optimize alternative connections between people and products [22].

**Figure 6: Client Layer**

Our proposal is an innovative architecture that supports a new model considering three fundamental elements:

*Numerous sources of information*, represented as a web service using a new technology called WSO2, specifically through the employment of the Data Services Server component which provides data from any information source by using web services. This means the logic application is independent from the information source no matter the format or technology in which the information is available.

*Logic Layer*, which contains Learning Analytics algorithms, consumes data provided by the previous layer and carries out the capture process, integration and storage of the information. Said information is stored in a dedicated repository on which Learning Analytics techniques are applied to generate reports and chart visualizations then available to the decision makers.

*The client layer*, presents the information in a proper manner, focusing on deploying relevant information through a HCI (Human Computer Interaction) mechanism to supply an easier interpretation to the end user, it also allows the use of this information in various forms.

As stated above, this innovative proposal relates different technologies to support decision making processes in an efficient manner, employing historic registers, information analysis and visualization systems, thereby freeing and involving different information sources, interrelating this data and presenting it in different visualization forms to support decision making.

4. RESULTS

Once the architecture has been developed through the RUP software development methodology (start, development, construction and transition) and deployed in a web application called UDLearn, the results obtained are shown in the following figures. As mentioned previously, the deployment of the web application was carried out in the Engineering Faculty of the Francisco José de Caldas District University. The graphics were designed taking into account the Human Computer Interaction principles: visibility, correct and clear feedback, restriction, mapping and matching, and consistency.

*Figure 7: UDLearn consult screen for student graduation documents.*
The above figure shows the user interface provided by the UDLearn web application and the three modules constructed which correspond with the objectives proposed: 1) Manage student grade documents, 2) Consult student academic achievement and 3) Consult Academic Practices.

On the top of that, three different graphs are shown in order to offer various perspectives of the information shown. In Figure 7, the participation level of the teachers in the students’ graduation work in each degree program is shown. In this particular case, 54% of teachers collaborated as directors of graduation studies and 36% as editors. Additionally a list below the graphs is shown summarizing the documents that each teacher was assigned and the kind of participation (director, editor or judge). The visualization of this information provides a tool to use when assigning new editors and judges considering the number of documents that each teacher has already been assigned. It is also useful to show higher academic administrators the participation level of teachers in collaborative academic activities.

Figure 8: UDLearn consult screen for interest level in graduation study topics.

Figure 8 enables the user to realize that 64% of the student graduation studies were focused on the topic of Software Engineering in the Systems Engineering program, and just 10% were focused on management and innovation technology. After conducting interviews with the curricular program administrator, and defining the investigation areas in each program and student graduation study, information such as student interest and job projection can be viewed through this consult screen.

On the other hand, looking at the consult screen in the student academic achievement module, queries for each student are observed, such as: average grade in each semester, average grade vs. student cohort average, grades in each subject, grades in different subjects assigned to teachers, subjects to be completed, and subjects approved, among others.

Figure 9: UDLearn consult scree for student achievement scale.

Figure 9 shows that a particular student has had 63 subjects in his academic program approved in contrast with the other students in his cohort who have had just 42 subjects approved on average. This means that the academic achievement scale of the student in question is 50% over the average. The visualization of this information helps to analyze and monitor the academic development of particular students and cohorts.

5. CONCLUSIONS

In this paper we have presented a vision of the system architecture of a Learning Analytics platform which can be used as a powerful
generic tool by Higher Education Administrators and educational decision makers, moreover its implementation shows the functionality of the model proposed and the various queries that can be managed through its functions.

Our proposal could be applied in different contexts other than education due to the fact that large volumes of information about diverse economic sectors are stored in multiple systems and technologies and this approach has considered this problem and proposes a mechanism to solve it.

Once the information that comes from the various repositories is available in an equal format (XML), it may be used in any way, in this case in a local repository; nevertheless information could be treated in other ways. The HCI visualization mechanism is a powerful tool that allows the interpretation of information from different perspectives, helping to facilitate better decision making by allowing diverse variables to be interrelated in order to analyze their correlation and impact on one another.

Our future work would be to implement this approach in the field of Artificial Intelligence to enhance the recommender system and use different data mining software as Weka or R, among others, to develop the data mining process that in the case of UDLearn was carried out during the ETL process and saved in modular XML queries to be used by the proposed modules.

REFERENCES


[18] “Data Services Server | WSO2 Inc.”.


