

Descriptive Analytics Dashboard for an Inclusive Learning Environment

Vladimir Costas-Jauregui <i>Centro MEMI</i> <i>Universidad Mayor de San Simón</i> Cochabamba, Bolivia v.costas@umss.edu.bo	Solomon Sunday Oyelere <i>Department of Computer Science</i> <i>Luleå University of Technology</i> Luleå, Sweden solomon.oyelere@ltu.se	Bernardo Caussin-Torrez <i>Ingeniería de Sistemas</i> <i>Universidad Mayor de San Simón</i> Cochabamba, Bolivia bernardocausin.t@fcyt.umss.edu.bo	Gabriel Barros-Gavilanes <i>School of Systems Engineering</i> <i>Universidad del Azuay</i> Cuenca, Ecuador gbarros@uazuay.edu.ec
Friday Joseph Agbo <i>School of Computing</i> <i>University of Eastern Finland</i> Joensuu, Finland friday.agbo@uef.fi	Tapani Toivonen <i>School of Computing</i> <i>University of Eastern Finland</i> Joensuu, Finland tapani.toivonen@uef.fi	Regina Motz <i>Facultad de Ingeniería</i> <i>Universidad de la República</i> Montevideo, Uruguay rmotz@fing.edu.uy	Juan Bernardo Tenesaca <i>School of Systems Engineering</i> <i>Universidad del Azuay</i> Cuenca, Ecuador juan.tenesaca@uazuay.edu.ec

Abstract—The educational community continuously seeks ways to improve the learner-centered learning process through new approaches like Learning analytics and its dashboard, which is helpful to enhance the teaching and the learning process. It involves a process whose final goal is presenting results to support decision-making about improving the learning process. However, a descriptive Learning analytics interface for analyzing learning data of students, including the disadvantaged, where to view and interpret learners' historical data is -in general- missing in this research domain. Hence, more research is still required to establish the philosophy of learning analytics on inclusion with an interface for the stakeholders to understand learning and teaching in an inclusive learning environment. This paper fills this gap by providing an inclusive educational learning analytics dashboard to support teachers and students. This study aimed to present a learning analytics implementation in the context of a smart ecosystem for learning and inclusion. We gave the inclusive educational needs and discussed the workflow followed during the descriptive learning analytics dashboard development. Therefore, the study improved existing learning analytics dashboards with a descriptive approach and inclusiveness of students with disabilities. Owing to the software development nature of this study, agile methodology based on five stages was applied: requirement elicitation; data gathering; design and prototyping; implementation; and testing and integration. We performed an initial evaluation, which indicated that the dashboard is suitable for understanding teachers' and students' needs and expectations. Besides, the visualization of inclusive learning characteristics improves engagement and attainment of learning goals.

Index Terms—descriptive learning analytics, inclusion, learning environment

I. INTRODUCTION

Learning analytics (LA) is helping to improve the teaching and the learning process. It involves identifying learning indicators, collecting learning data, analyzing the collected data, and presenting results to support decision-making about improving the learning process. Nowadays, there is much emphasis on a learner-centered approach. The educational

community continuously seeks ways to improve the learning process through new strategies, technology, and content [1,2].

Descriptive analytics is one of the most popular approaches to Learning analytics development [3]. Learning analytics uses log-data and assessment data extracted from the Learning Management System (LMS) and other sources for further targeted analysis. Statistical methods such as descriptive statistics remain helpful in understanding the interplay between various learning indicators and are applied to obtain results that help make decisions [4,5]; this case regarding learning progress by both the teachers and learners. Descriptive statistics' main goal is to understand the data by analyzing the summary features of large amounts of data [6].

The dashboard must provide helpful information to allow the student to visualize learning progress and identify areas that will need some support to achieve the learning objectives [7]. A Dashboard is a suitable one-screen visualization of learning analytics results for presenting to the student the awareness of the impact of their efforts. Understanding the implications of the effort is essential for students to sustain motivation and focus [8]. Moreover, the dashboard helps the teacher to reflect and improve on their teaching practice.

Learning analytics to be used by teachers and students -data science novices- must be intuitive, simple, and easy to understand. The systems implemented in the past have suffered in flaws where the learning curve is steep [9–12]. The simplicity and power of the dashboard is an ongoing problem by the visualization choice and user understanding [13,14]. There are findings of how hard it is to understand visualization results by the students and teachers because of insufficient data literacy (data reduction, dimensionality reduction) [3]. Reference [15] states that users' shared understanding of data and visualization is for extracting, integrating, and finding a relationship in the data from primary to middle-level data visualization comprehension. On the other hand, bar charts are a required visualization by the users because of their simplicity.

These studies show the user has no good understanding of visualization; they prefer simple data representation like a bar chart and pie chart. But, researchers require descriptive information to find and analyze implicit relations [15].

However, research has been sparse about how to use learning analytics methods to support inclusive education. Previous research has shown that many students have specific learning disabilities in the current educational system [16]. According to [17], an estimated 34% of students in the US between ages 3 and 21 have particular learning disabilities. Many students with learning disabilities lack access to information about their learning progress. Reference [18] found that students with disabilities are not engaged in the learning process, which ultimately affects their learning progress. Because of the apparent shortcomings of the current learning analytics dashboards, studies do not support students with disabilities.

This study will present how to build from scratch, using reusable tools, a visualization dashboard suitable for supporting students with learning disabilities, and if this kind of dashboard visualizing inclusive features is helpful to improve course progress and assessment to students with learning disabilities. The design and development of the learning analytics dashboard start and progress together with the Smart Ecosystem for learning and inclusion (SELI) platform, as the agile methodology drives the development of the platform, the agile approach based on five stages was applied: requirement elicitation; data gathering; design and prototyping; implementation; and testing and integration. An open interview with a teacher gives feedback on dashboards' helpfulness. We present the inclusion educational needs and issues for which to get answers with learning analytics. Finally, we discuss the workflow followed while developing a descriptive learning analytics dashboard with reusable tools and the feedback from the open interview.

II. LITERATURE REVIEW

Descriptive learning analytics uses educational data to understand the past and the present. Descriptive learning analytics does not predict future events but instead puts historical data into context through visualizations and other interpretable methods. Such methods include comparative methods such as presenting mean and median values of the features within the dataset. Predictive and prescriptive learning analytics, in general, use more complex data modeling than descriptive learning analytics [19,20], but understanding the historical and current data is crucial to many learning contexts.

Common methods to model the historical data are charts (visualizations) and tables (comparative) and are commonly presented in dashboards available for students, teachers, and other stakeholders. Multiple dashboards-based learning analytics tools have been developed over the years. They include LOCO-Analyst, the Student success system, and SNAPP, which all are designed to be used by the teachers. Course-signal and Narcissus are dashboard tools intended to be used by students. Student Inspector, GLASS, SAM, and StepUp! are dashboard-based descriptive learning analytics tools developed

for both teachers and students [21,22]. All of these tools involve descriptive Learning analytics elements, but Course-signal also includes predictive algorithms.

Descriptive learning analytics can be used for formal or informal purposes depending on the raw data context and the proposed stakeholders. Whether the system is designed to be formal or informal, the aim is to present an interpretable and justified presentation of raw data collected from the learning environment, such as the log files of a digital learning environment. Dashboard-based descriptive learning analytics can influence the teachers' pedagogical dynamics to fit students in different contexts [23]. Educational institutions are leveraging the potential of dashboard-based learning analytics to understand how students interact and derive knowledge from various components provided within learning platforms.

Learning analytics dashboard can be teachers-centered, students-centered, or a hybrid depending on the contextual need. For example, Salcik in [24] conducted a teacher-centered learning analytics dashboard of a visualized learners' behavior and actions in an educational system by tracking their interaction with widgets in the application. Similarly, [25] studied how instructors used learning analytics in an educational platform. Instructors' historical data of using the platform, including how often they visited the Learning analytics dashboard, duration of use, and the dashboard component used, formed part of the data analyzed [26]. On the other hand, to make learning analytics valuable for students, [26] explored the design and implementation of a student-centered learning analytic dashboard in the higher education context. According to [26], the involvement of students in the development of dashboard-based learning analytics provides them with some level of control over learning analytics, which can significantly improve their self-regulated learning and academic performance. Besides, students can see their learning history in a visualized learning dashboard, positively affecting their attitude towards learning.

Reference [27] revealed that the development of Learning analytics dashboards meant for students' and teachers' use is growing. However, learning analytics dashboards for inclusive educational applications that collect learners' data while learning to provide descriptive analysis is not concretely researched. Besides, inclusive educational platforms that provide descriptive learning analytics are not a widely studied research area. Hence, more research is still required to establish the philosophy of learning analytics on inclusion by providing an interface for the stakeholders to understand different types of learning and teaching in an inclusive learning environment. In other words, a descriptive learning analytic interface for analyzing learning data of students, including the disadvantaged, where learners' historical data can be viewed and interpreted is -in general- missing in the descriptive learning analysis platforms. This paper fills this gap by providing an inclusive educational learning analytics dashboard to support teachers and students.

III. DESIGN AND DEVELOPMENT PROCESS

The agile approach has been used in the software development process to facilitate the rapid development of a learning analytics tool with the Smart Ecosystem for Learning and Inclusion (SELI) platform. Agile software development processes were designed primarily to address the problem of timely deployment of software to customers. According to [28], the following four attributes are helpful to define an agile approach: incremental, cooperative, straightforward, and adaptive. The same four attributes appear in the six attributes that suit an analytics development style [29]; evidence of the agile approach benefits in building learning analytics tools in the LALA project [30]. The main benefits are: tiny software releases in rapid cycles, developers work collaboratively with the stakeholders throughout the life cycle of the process, learnability and adaptability of the method, and ensure good documentation.. The adaptive attribute allows for flexibility to make changes at the last moment. During the process of developing the learning analytics tool, the SELI developers and the research team adapt and adjust the changes derived from experiences gained during the development, changes in software requirement, and changes in the development environment to make minor releases of versions of the learning analytics tool to conform with the incremental attribute of the agile approach in exploratory projects and small development teams [31]. Several stakeholders within the SELI project, such as teachers, students, developers, regional partners, and project coordinators, were collaboratively engaged through communication and interaction to define the tool's requirement and subsequent development activities. The design and developmental effort of the tool were well documented, including providing comments within the software code. In the case of the adaptive attribute, the development process allows stakeholders to suggest last-minute changes, which the developers immediately implemented.

A. Indicators Stage

The main goal of elicitation is to collect the indicators as requirements from stakeholders. The Learning analytics indicators express what we want to know about participants' behavior in the learning process; with the indicators, an intervention could occur during learning.

For the SELI Learning analytics component, we identify two stakeholders: the teacher and the student. During the elicitation, the team inquired about the requirements by weekly meetings and workshops with the stakeholders. In the first meetings, the initial indicators surge through brainstorming; in the following discussions, the team and stakeholders tune the indicators. The participants discuss prototypes describing the visualization of indicators in the SELI learning analytics component.

The workshops take place in distinct countries. The team meets stakeholders to know their requirements and perceptions from the platform prototype; it includes the insights for intervention during the learning process. These insights let to collect the general inclusion issues warned by teachers

and students in a learning platform. After the workshops, the researcher and developers matched the initial indicators proposed against the ones perceived during workshops.

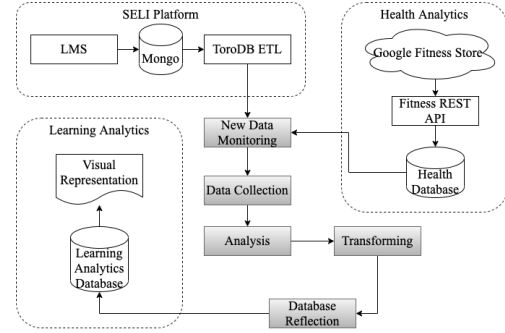


Fig. 1. Data transformation Workflow LMS to Learning and Health Analytics.

During the elicitation process, we identify the following indicators related to inclusion from the teacher: Number of courses with special requirements support, Relation of accessibility support in resources against the total of resources, Relation in a course with accessibility requirement support against accessibility support in resources activities, Match relation from each accessibility support against accessibility requirement, Students accessibility support usage, and Students performance against accessibility support use. Indicators related to inclusion from the student: Student resources preference about accessibility(to be implemented in next version), Course progress against student plan, Number of activities completed against student plan, Relation of activities completed successfully against failed, and Health goals related with the study plan.

For the administration role, the following main indicators list is a product of the elicitation process: Number of teachers and students enrolled in the platform, Number of interactions between students and teachers based on reading and writing actions in the courses, Time access log to courses (teacher/student), Number of users, Number of total activities of all courses, and How the student is moving on between courses (the path or trajectory). The Learning Analytics component will compute indicators based on available data, in this case from the SELI platform and an external service related to health data for students.

B. Data Gathering Stage and Establish the Requirements

The SELI Learning Management System feeds a non-relational database; in the next step, the ToroDB¹ tool read from MongoDB to extract, transform and load the data into a relational database (PostgreSQL). Finally, the building of visualization in the SELI dashboard is a service from the Metabase² tool. Fig. 1 shows the flow from data to analytics and visualizations dashboard process.

¹ToroDB is an application to replicate and transform a MongoDB into PostgreSQL database. <https://www.torodb.com/>

²Metabase is an open source analytic tool. <https://www.metabase.com/>

For the student learning analytics, the execution and acquisition process in the Data-Analyze-Plan-Execution monitoring-Reflect model (DAPER) from [32] is adopted, following the five steps shown in gray in Fig. 1. For this student analytics supporting health and learning goals, there is extra health data gathered from Google Fit. The health data and the student's health goals let the analytic engine compute the percentage of accomplishment and prediction model about values. The reflection will suggest to the student how to improve or keep going with her health goals. The help goals suggestions are: complete your steps goal, your need to sleep more, congratulation you reach your calories goal. In the same way, the student provides goals for his learning about the task/activities completed during the course, the time to spend in a course, the weekly plan of progress in the course; at the end, the student will perceive her progress from the following four possible suggestions from the prediction model: please go for your incomplete task, please keep pace to complete the course, this plan needs modification to reach the goal, complete your work.

The student health goals are the standard Google Fit data related to counting, sleep time, and daily calories burn. Google Fit will collect these values from a device like a mobile phone or other wearables. The students' learning goals about weekly activity/task to complete, weekly course peace plan, planned time to complete, and the deadline for tasks related to data in the SELI LMS database; the analytic engine will gather the values compute results to compare against goals. The student will provide the values for health and learning goals. The data monitoring trusts in the data extracted and transformed from the SELI platform and external data related to health (Google Fit service). The next phase is the compatibilization of data as a unique data collection to feed the analysis process against goals. Finally, the analyzed data transformation results in a new reflective database for future analysis.

1) *Raw data*: The database stores Spatio-temporal data on the student's interaction with the teacher's activities, such as questionnaire homework and storytelling; activities with the interaction between students and teacher chat and forum have data stored in the database. Similarly, Spatio-temporal data related to access to resources like video, audio, a text document, and file. For each activity, data such as delivery time, files or content editing, and student performance (teacher feedback). Also, event-driven data allows leaving a trace of the student in authentication and navigation across the different aspects of the learning environment.

2) *Indicators*: The indicators (described in III-A) allow us to answer questions by looking at the data using patterns and calculations quantifying the observations. To deliver learning analytics visualizations that will aid decision-making, we need to process the data. However, the "raw" data are stored in a non-relational database because MongoDB is part of the Meteor framework used for the SELI Platform deployment and faster development. This non-relational database must be preprocessed to obtain a dataset suitable for the analysis to support the teacher's decision-making process. Transforming from the non-relational to a relational model is an essential

preprocessing step to an effective and efficient query process. Moving to a relational model implies data fitting towards maintenance and an appropriate stable data structure model (i.e., some data have more or fewer attributes than others of the same type). Therefore, it is necessary to obtain all the possible attributes with corresponding values and assign a null to missing attributes value; it is a standard solution for missing values. The missing values in MongoDB are the result of an evolving data model along with the development and maturity of the platform, and null during analytics do not participate in statistics nor queries formula to compute data results. With the data represented in a relational model, it is possible to make queries, allowing the analysis to obtain descriptive statistics (mean, variance, standard deviation) suitable for a basic understanding of learning progress by both the teacher and students. The non-relational data allow us to obtain comparisons letting the teacher make decisions.

MongoDB has a functionality, a replica set, which provides a database replica that ToroDB uses to process without damaging the original set. ToroDB is an ETL (Extract, Transform and Load) used to extract MongoDB data, transform it into SQL (Standard Query Language) and load it into a PostgreSQL database. The virtue of this tool allows the synchronization between the relational and non-relational databases in real-time. The synchronization follows a couple of rules: MongoDB fields are transformed into columns of the same name, with a suffix that indicates the type of data and the interpretation that ToroDB gives to that column. The `Null` means the absence of data in the MongoDB documents. This kind of value also affects the suffixes that ToroDB adds to the columns' names; for example, the suffix "_n" means a `null` value, stored with PostgreSQL type boolean (nullable). It cannot take value false, just true or null. When the value is true, the JSON document has null for that path; when it is null, it means the path has another value or does not exist for that document. Keys are transformed into new tables in Postgres, using primary id.

3) *Data processing*: Metabase access and manage the PostgreSQL database for the analytics. Metabase is a web interface for creating indicators through SQL and allows inferred knowledge visualization. Additionally, Metabase provides many ways to visualize the assumed knowledge in real-time forms of charts, tables, maps, lines, and combo. The inferred knowledge is the information that the teacher and student see on the dashboard. Metabase queries can be simple or can include parameters (ID of student or teacher, configuration). This delimitation by parameters helps to constrain data in the visualization and use recycled queries (i.e., data only related to a particular user). These parameters are only accessible to the administration user, ensuring user data anonymization and additional user data restrictions.

4) *Visualization*: It is important to view each type of user on the SELI platform: students, teachers, and administrators. All of these users have specific indicators, which are visualized through the dashboard. Metabase allows the creation of panels on the dashboard in a personalized way. It groups several queries in the same place, allowing to see globally or,

precisely, the required indicators. The teacher can view these queries through embedded Metabase views, and these views can receive parameters to distinguish the graphs depending on the user who consults them. For example, the teacher dashboard shows all the indicators of a selected course; they can download each chart and review their course's general behavior or other selected courses.

Learner's lifestyle can facilitate self-directed learning. The study has shown that students engage in other activities beyond the academic activities that can affect their learning [32]. One way to ensure a healthy learning lifestyle is to seek the advice of professional experts. However, a computer-based intelligent system can improve self-direction skills (SDS) with an automated dashboard [33].

The students learning analytic dashboard was developed for self-control to regulate academic and nonacademic skills based on the DAPER model [32]. The application collects users' learning and health data, visualizes them inside the dashboard, and analyzes to generate a plan to improve a healthy learning lifestyle. The learner can also monitor personal actions and decide if any re-plan is necessary to achieve the desired goal.

In the administrative dashboard stands the general status of the platform. It let the administrator check the following: the number of students and teachers registered, the available interaction on the platform, and the users' initiation of sessions.

On the teacher dashboard, you can see the number of student interactions in the teachers' courses, separated by date and compared to the average. In Fig. 2, the purple line lets us know the average interaction along with course duration. The teacher can see the variation between laziness and interaction peaks; it will permit actions to reduce the distance from the ground to peaks according to due dates and time planned per week. The teacher notices when the students carry out activities or watch resources; thus, the visualization shows the preference daytime of the students to access the course. Suppose the students have a high preference to access in the afternoon, and the morning is the lazy time. In that case, the teacher should avoid activities finishing at noon and take into account fewer students are enabled to work during the morning.

The SELI platform lets the student warn about her disabilities. The system requests an accessibility feature when the course subscription; in the case of motor disability, the student will need extra time or limitless time to work on a quiz.

Usually, teachers provide video-recorded classes; in the SELI Platform, there is the option to attach accessibility features like sign-deaf language, captions, and text alternatives for inclusive learning resources. On the teachers' Dashboard, you can see Fig. 3, a visualization related to video and accessibility commonly used in the course. The percentage of use helps to plan future courses and improve the availability of resources supporting inclusive needs. The graphic in Fig. 4 complements the information about video accessibility. In this chart, there is a comparison of accessibility between students requiring accessibility options, probably because of a disability, and students not demanding accessibility features. Fig. 5 shows the student video accessibility demand compared to accessibility

alternatives provided in the video resources. The chart stands a low deaf-sign language availability against students' demand.

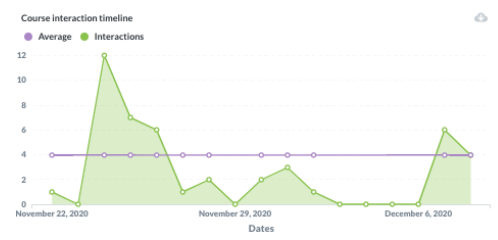


Fig. 2. Course-Student interaction timeline.

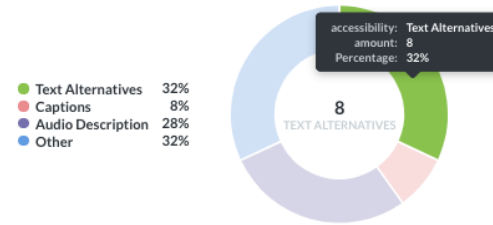


Fig. 3. Video Accessibility available in the course.

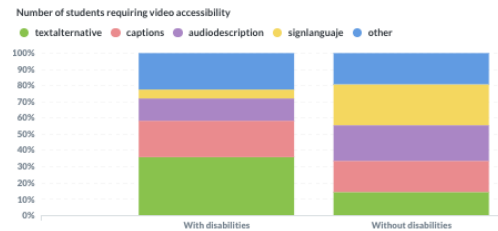


Fig. 4. Student video accessibility requirement in the course.

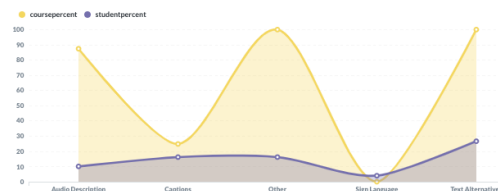


Fig. 5. Student video accessibility requirement against provided.

5) *Security*: The learning analytics component has two service components: ToroDB and Metabase. The former deals with ETL as a service consuming from SELI database; the latter deals with analytics and provides a dashboard. Both of them have an administrative user setup in the SELI server; this user helps set up and look for logs to ensure everything is working well. Both components and the SELI Database live in the same server and trust in the Operating system's security setup, MongoDB, PostgreSQL, and Meteor framework. The ToroDB service uses a MongoDB user to extract data from the SELI database and write it up to a PostgreSQL database. The Metabase Dashboard service exposes through an HTTPS

connection and encrypts database detail in REST messages. Metabase reads from the PostgreSQL database to make the analysis. Security between MongoDB and PostgreSQL replica uses standard ODBC security communication.

The dashboard visualizations are accessible from any web page using a security token. The tokens and additional parameters travel through HTTPS requests in the same embedded views of each user, guaranteeing the confidentiality of the user's data in the SELI learning platform.

C. Design/Prototyping

The SELI Learning Management System is developed using new technologies like Javascript language through Node.JS and MongoDB. These technologies could be considered new compared to environments used for well-known Learning Management System open source software from the last 20 years (e.g., Moodle). Special emphasis is placed on frameworks allowing component-oriented and fast prototyping of solutions (i.e., Meteor).

Because of this fast-paced development, a quick solution for providing a proof-of-concept was required. Additional requirements include integrating existing, stable open-source projects, mainly because of the benefits of code reuse. To the best of our knowledge, it is the first solution using Metabase for LA. However, these separated developments lead to different programming languages and types of databases. The platform uses Node.js and MongoDB, while the Learning Analytics module is developed using Java technology and PostgreSQL.

At the moment, no solution for graphic query builder is implemented on the dashboard. The main reason is that understanding the internal structure of data is a time-consuming task for most teachers. Our approach requires a specialist or a highly proficient user to generate indicators and make them available to end Learning analytics users.

In general, the system obtains analytics: for the learner, for the educator, for curriculum responsible, and inclusion. This proof-of-concept implementation generates indicators related to general Learning Analytics and specific indicators about inclusion through graphs and charts.

D. Testing and Integration

The nature of the system development leads to having the Learning Analytics module advancing faster than the Learning Management System. Periodically, a database snapshot from the system was shared with developers from the Learning Analytics module, just simple structured data without multimedia records like images or videos.

Two independent mirror servers were used in the development, one for the Learning Management System and one for the Learning Analytics module³. Nevertheless, some integration problems arose when configuring the software of two servers in just one machine. In any case, Metabase-generated information could be accessed by logging in as a tutor in the Learning Management System. To achieve this, it was

³Actually, both servers are in the same equipment, the LMS and LA component.

necessary to turn off the Nginx service used in the container of the Learning Management System deployment and make a manual configuration of Nginx directly in the machine. This redirects the consults to the respective domain to hold them in one server.

IV. EVALUATION FROM A TEACHER THROUGH INTERVIEW

This current study is yet to conduct a detailed evaluation with teachers and students who are the targeted users of the learning analytics dashboard but presented reflective feedback from a teacher who conducted an online course on the SELI platform. Although plans for a comprehensive evaluation of the developed learning analytics dashboard are ongoing, this section presents the content analysis of a teacher's experience gathered through an interview. The authors intended to know whether the dashboard visualization of students learning analytics enhances the teacher to track students' learning progress in a course. The response shows positive as the teacher asserts

"Yes, you can have a general course progress average and the percentage of progress for each student."

Besides, the teacher also acknowledged that the students learning progress could be monitored individually by the percentage provided on the dashboard. To some extent, the dashboard allows for monitoring of students' interactions and their lazy periods.

Regarding the opportunity to interpret inclusive use of resources and activities by the student during course progress through the visualized dashboard, the teacher agreed that some elements of this feature were seen. For example, a response to this question reads

"you can see how many students required extra time for a quiz; You wonder that half of the students claimed for it. It means that they have a need, but we do not have this info for all the student's inclusive claims; part of the visualization stands the requirements from the student against the teacher's inclusive features"

Furthermore, the teacher provided a perspective of what could be added to the current learning analytics dashboard to improve users' experience. One of the features the teacher wishes to see includes the analysis of how students used the accessibility features provided by the teacher while creating the course. According to the teacher, graphical information about students' progress with a disability would help evaluate how inclusive their learning and teaching have been over a period. Specifically, the teacher asserts

"I think it is important to add knowledge about the interaction with the accessibility provided by the teacher and understand if the student uses this kind of accessibility; the student fills out her needs about inclusion, and the teacher provides some accessible features and are these features used by the needed students? In what percentage? This kind of question we need to answer and try to figure out the kind of features to improve"

Finding out whether the dashboard is easy and straightforward to manage, the teacher gave feedback that suggests a need to improve the graphical representation and readability of

the visual representations and visualizations to enhance better comprehension. In this sense, suggestions from the teacher read

“probably improve the visualization title or add some extra description of what stands in this figure.”

These suggestions provide helpful feedback to improve the solution before making large-scale experimentation where teachers and students participate in the study.

V. ETHICS AND DATA PRIVACY

The success of Learning analytics depends on the quality and quantity of data it collects. However, always consider an ethical perspective about the collection and use of private data. The Learning analytics community has been examining the impact of ethics and privacy [34,35], and the Open University in [36] has established eight essential principles for the ethical use of student data to carry out Learning analytics.

Reference [37] described several Learning analytics scenarios and identified principles where ethical issues arise. In March 2019, the International Council for Open and Distance Education (ICDE) published a report with guidelines for applying ethically informed Learning analytics [38]. This report highlights ten points that we will enumerate and discuss how SELI implements it:

1) *Transparency*: Institutions must make the purpose of Learning analytics clear to students and other stakeholders. The SELI project achieves this by conducting open workshops to prepare teachers for the use of Learning analytics. During three face-to-face workshops offered by SELI Team in Uruguay (at Montevideo, Rivera, and Melo in February 2020), and the course Smart Ecosystem for Learning in SELI platform with Learning analytics as one of their central subjects. SELI Learning analytics proof-of-concept provides access to a detailed specification of the analytical data process from an infrastructure perspective.

2) *Ownership and control of data*: In our case, the national legislations among Europe, Latin America, and the Caribbean are distinct; it complicates the implementations of data privacy regulated by national and international legislation. For example, there is a lack of clarity about who owns the data (institutions versus students). As an interoperable platform among countries on different continents, the SELI future implementation would accomplish a strict regulatory framework, where limitations and principles are respected when we collect, process, store, and analyze personal data, following the recommendations of [39–41]. The institutions establish the temporality of the data collected and consider the students’ data as sensitive and personal.

The SELI platform has not developed mechanisms for these features yet. But we are exploring the design as an ecosystem made up of interactions between services, some emerging technologies such as Blockchain, IPFS (InterPlanetary File System), and Smart contracts as strong candidates to ensure ownership and control of data and as a promise to get significant advantages [42–45]. On the other hand, works by [46–48] show limitations of blockchain to manage data privacy.

Blockchain currently cannot solve the data privacy problem since, by design, any connected node can see the information uploaded by users. However, Smart contracts offer a way to secure access to the data; together, data protection mechanisms such as encryption result in a secure and reliable system.

3) *Accessibility of data*: It relates to determining who has access to raw and analyzed data and clarifies which data would be inside a Learning analytics application and what data might be out of scope (i.e., income, academic history, demographic data, etc.). In the SELI project, the accessibility of data relates to the capacity to manage data despite the form of their representation, with an effort to get the Learning analytics accessible for different capabilities, i.e., textual, visual, etc. We are working to complement this accessibility of data point of view to ensure accessibility using a peer-to-peer architecture with IPFS protocol as currently being proposed by [49,50].

4) *Validity and reliability of data*: Refers to ensure the data collected and analyzed is accurate and representative of the problem. It is a sensitive-essential issue for predictive calculations (the SELI project only implements descriptive analytics at the moment). However, the data collected for description is accurate and representative, but we had identified some distortion for future predictive calculations; for example, a student can enroll twice in a course on the platform. We need more tests on the platform and discover this kind of issue.

5) *Institutional responsibility and obligation to act*: It raises the importance that institutions reflect on whether access to knowing and understanding more about how students learn generates a moral obligation to act. The institution must consider its policy to identify how it will affect the data obtained with the Learning analytics application. In this sense, for future work on SELI, we plan to provide a particular interface for the entry of interventions and their recording of impacts, although the possibility of having these functionalities are not typical characteristics in current LA systems.

6) *Communications*: The guideline points out the sensitive characteristic of communications when communicating directly with students based on their analytics. We believe that this topic highlights the interdisciplinary character that must exist in LA projects. Communications in a student-centered learning analytic system cannot be messages written by software developers; instead, we plan to be a redaction team of teachers, communication experts, and psychologists.

7) *Cultural values*: The subject of Cultural values indicated in the guideline is very relevant to the philosophy of the SELI project, which defends learning in diversity. In multicultural contexts, with migration and refugee problems, the understanding and interpretation of data are necessarily more complex. It is necessary to develop a Learning analytic application from an interdisciplinary approach integrating (in the team) disciplines like Sociology and Anthropology.

8) *Inclusion*: The management of Inclusion from the ethical perspective can assume greater risks if the institution uses Learning analytic to improve its categorization in different rankings. There is a risk of using Learning analytic to legitimize the exclusion of certain students instead of help

them. The Learning analytic would collect the actions followed by the institution and make them available for regulatory organization evaluations. The SELI Learning analytics system does not implement this functionality.

9) *Consent*: The consent request should be on registration, expressing an explicit description of the personal data collected, detailing the purpose of use, data temporality, security policy, and the person responsible for the data processing [40]. But the students may not completely understand this amount of information. They may be unable to evaluate the implication of using their data at the time of registration in their correct dimension. For this reason, a suggested approach is to follow the criteria "privacy by design" [51]. The approach includes all data privacy issues among the initial system requirements and includes lawyers in the interdisciplinary team [52].

10) *Student Responsibility*: As established in [53], the Learning analytic application shows the power relations institutions and teachers have over students. To mitigate this scenario, the institutions should ensure the students' involvement in the creation and design interventions that will support them. It is a future work to improve the SELI Learning analytic system.

VI. CONCLUSIONS

The intention to present an improved learning analytic dashboard with inclusive information, in a descriptive way, helps to decide how to improve students' accessibility. The inclusive features strengthen the learning process in inclusive education.

An agile development process is the right choice for the SELI platform, concretely for the Learning analytics module. The development of the authoring tool and course area where the student works is incremental. For each feature implemented, the learning analytic development team must start the event incrementally catching in the platform, and the dashboard visualization related to the indicator identified. The success is the cooperative work between both development teams, developers, and research teams. The research team warns about accessibility issues from universal learning design testing; this warns changes, sometimes significant, in the platform's components. They affect the data collected for Learning analytics; the team stated a good adaptive reaction. The group learned and adapted the process from scratch since the first stage, keeping track of the process with a well-documented tool. As stated in other experiences, our process's weakness is the decentralized teams working in distinct SELI platform modules.

Using a tool like Metabase helps rapid development; but, you have some restrictions because of the kind of visualization and the navigation way of Metabase. In the analytic learning dashboard for students, the building of visualizations is from scratch with Charts.js, a JavaScript library; this approach shows different visualization than Metabase, but with extra time for development.

The SELI platform is still in beta and evolving; the Learning Analytics component need to add visualization suggested by

teachers and students who used the dashboard. Increment features sometimes mean updating the course area's events, where students interact with the course's resources and activities. The visualizations selected fit some needs of teachers according to the interview with one teacher. They also help them to improve the course and follow the progress of students with different abilities. The input is informal with one teacher, and some comments about the student dashboard from the research team and students from workshops carried out in the SELI project. From this feedback, we conclude that the dashboard helps the teacher improve the accessibility of the resource and let him know when the student works and which needs the course is covering. Similar findings are for the student dashboard, the research team tested the goals against progress in a class, and they believe the student will have helpful feedback about health and progress.

Future work will focus on evaluation and validation of the tool; we plan to qualitatively measure the impact of analytics results in improving a course and the impact on the students' progress. We will also explore different tools for visualization to solve restrictions in the visualizations and navigation issues found with Metabase. Another area to look for strengths is the agile approach in the development of the learning analytics tool. In this experiment, the agile methodology worked well together with the analytics cycle. However, a formal framework on how it works as twisted software development and analytics process that suits principles and methods needs more study.

ACKNOWLEDGMENT

This work was supported by the ERANET-LAC project, which has received funding from the European Union's Seventh Framework Programme. Project Smart Ecosystem for Learning and Inclusion – ERANet17/ICT-0076SELI.

REFERENCES

- [1] A. M. Volungevičienė, J. M. Duart, J. M. Naujokaitienė, G. M. Tamoliūnė, and R. M. Misiulienė, "Learning Analytics: Learning to Think and Make Decisions," *The Journal of Educators Online*, vol. 16, 2019.
- [2] J. Yanchinda, P. Yodmongkol, and N. Chakpitak, "Measurement of Learning Process by Semantic Annotation Technique on Bloom's Taxonomy Vocabulary," *International Education Studies*, vol. 9, no. 1, p. 107, 2015.
- [3] X. Du, J. Yang, B. E. Shelton, J.-L. Hung, and M. Zhang, "A systematic meta-Review and analysis of learning analytics research," *Behaviour & Information Technology*, vol. 40, no. 1, pp. 49–62, 2019.
- [4] M. Fleckenstein and L. Fellows, "Data Analytics," *Modern Data Strategy*, pp. 133–142, 2018.
- [5] U. R. Hodeghatta and U. Nayak, "Introduction to descriptive analytics," *Business Analytics Using R - A Practical Approach*, pp. 59–89, 2017.
- [6] R. Donnelly and F. Abdel-Raouf, "Let's get started," in *Statistics*, Third ed., Alpha, 2016. ISBN 978-1465451668
- [7] S. Few, "Dashboard Confusion Revisited," *Perceptual Edge*, Mar-2007. [Online]. Available: http://mail.perceptualedge.com/articles/visual_business_intelligence/dboard_confusion_revisited.pdf. [Accessed: 20-May-2019].
- [8] S. Charleer, J. Klerkx, E. Duval, T. De Laet, and K. Verbert, "Creating Effective Learning Analytics Dashboards: Lessons Learnt," *Adaptive and Adaptable Learning*, pp. 42–56, 2016.
- [9] T. Toivonen, I. Jormanainen, and M. Tukiainen, "Augmented intelligence in educational data mining," *Smart Learning Environments*, vol. 6, no. 1, 2019.

- [10] S. Kausar, S. S. Oyelere, Y. K. Salal, S. Hussain, M. A. Cifci, S. Hilcenko, M. S. Iqbal, Z. Wenhao, and X. Huahu, "Mining Smart Learning Analytics Data Using Ensemble Classifiers," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 12, p. 81, 2020.
- [11] K. Sunday, P. Ocheja, S. Hussain, S. S. Oyelere, B. O. Samson, and F. J. Agbo, "Analyzing Student Performance in Programming Education Using Classification Techniques," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 02, p. 127, 2020.
- [12] O. S. Balogun, S. S. Oyelere, and D. D. Atsa'am, "Data Analytics on Performance of Computing Students", New York, NY, USA, 2019.
- [13] W. W. Eckerson, *Performance dashboards measuring, monitoring, and managing your business*. Hoboken, NJ: Wiley, 2011.
- [14] M. Elias, M.-A. Aufaure, and A. Bezerianos, "Storytelling in Visual Analytics Tools for Business Intelligence," *Human-Computer Interaction – INTERACT 2013*, pp. 280–297, 2013.
- [15] A. Sarikaya, M. Correll, L. Bartram, M. Tory, and D. Fisher, "What do we talk about when we talk about dashboards?," *IEEE transactions on visualization and computer graphics*, vol. 25, no. 1, pp. 682–692, 2018.
- [16] Lukasz Tomczyk, S. Sunday Oyelere, and others, "ICT for Learning and Inclusion in Latin America and Europe. Case Study From Countries: Bolivia, Brazil, Cuba, Dominican Republic, Ecuador, Finland, Poland, Turkey, Uruguay", 2019.
- [17] U.S. DEPARTMENT OF EDUCATION, W. J. Hussar, and T. Nachazel, Washington: U.S. DEPARTMENT OF EDUCATION, 2020.
- [18] E. Foster and R. Siddle, "The effectiveness of learning analytics for identifying at-risk students in higher education," *Assessment & Evaluation in Higher Education*, vol. 45, no. 6, pp. 842–854, 2019.
- [19] P. Arroway, G. Morgan, M. O'Keefe, and R. Yanosky, "Learning analytics in higher education", Research report. Louisville, CO: ECAR, March 2016. 2016 EDUCAUSE. CC by-nc-nd, 2015.
- [20] V. L. Uskov, J. P. Bakken, A. Byerly, and A. Shah, "Machine Learning-based Predictive Analytics of Student Academic Performance in STEM Education," 2019 IEEE Global Engineering Education Conference (EDUCON), 2019.
- [21] A. Gruzd and N. Conroy, "Learning Analytics Dashboard for Teaching with Twitter", 2020.
- [22] M. Salioun, "State of Art of Data Mining and Learning Analytics Tools in Higher Education," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 21, p. 58, 2020.
- [23] I. Molenaar and C. Knoop-van Campen, "Teacher Dashboards in Practice: Usage and Impact," *Data Driven Approaches in Digital Education*, pp. 125–138, 2017.
- [24] S. Salkic, S. Softic, B. Taraghi, and M. Ebner, "Linked data driven visual analytics for tracking learners in a PLE", *DeLFI 2015–Die 13. E-Learning Fachtagung Informatik*, 2015.
- [25] S. L. Dazo, N. R. Stepanek, A. Chauhan, and B. Dorn, "Examining instructor use of learning analytics", in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2017, pp. 2504–2510.
- [26] L. D. Roberts, J. A. Howell, and K. Seaman, "Give Me a Customizable Dashboard: Personalized Learning Analytics Dashboards in Higher Education," *Technology, Knowledge and Learning*, vol. 22, no. 3, pp. 317–333, 2017.
- [27] B. A. Schwendimann, M. J. Rodriguez-Triana, A. Vozniuk, L. P. Prieto, M. S. Boroujeni, A. Holzer, D. Gillet, and P. Dillenbourg, "Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research," *IEEE Transactions on Learning Technologies*, vol. 10, no. 1, pp. 30–41, 2017.
- [28] P. Abrahamsson, N. Oza, and M. T. Siponen, "Agile Software Development Methods: A Comparative Review1," *Agile Software Development*, pp. 31–59, 2010.
- [29] "Agile Analytics Framework Overview and Best Practices," XenonStack, 08-Dec-2018. [Online]. Available: <https://www.xenonstack.com/insights/what-is-agile-analytics/>. [Accessed: 22-Jun-2021].
- [30] H. Chevreux, V. Henríquez, J. Guerra, and E. Scheihing, "Agile development of learning analytics tools in a rigid environment like a university: Benefits, challenges and strategies", in *European Conference on Technology Enhanced Learning*, 2019, pp. 705–708.
- [31] N. Keshta and Y. Morgan, "Comparison between traditional plan-based and agile software processes according to team size & project domain (A systematic literature review)," 2017 8th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2017.
- [32] R. Majumdar, Y. Y. Yang, H. Li, G. Akçapinar, B. Flanagan, and H. Ogata, "GOAL: Supporting learner's development of self-direction skills using health and learning data", in *Proceedings of the 26th International Conference on Computers in Education (ICCE2018)*, pp. 406, 2018.
- [33] M. J. Rahman, "Self-Regulated Learning Platform Based on Learning and Health Data", Master's Thesis, Itä-Suomen yliopisto, 2020.
- [34] D. Ifenthaler and M. W. Tracey, "Exploring the relationship of ethics and privacy in learning analytics and design: implications for the field of educational technology", *Educational Technology Research and Development*, vol. 64, no. 5, pp. 877–880, 2016.
- [35] D. Gasevic, S. Dawson, and J. Jovanovic, "Ethics and privacy as enablers of learning analytics", *Journal of learning Analytics*, vol. 3, no. 1, pp. 1–4, 2016.
- [36] The Open University, Ed., *Policy on Ethical use of Student Data for Learning Analytics*, Sep-2014. [Online]. Available: <https://help.open.ac.uk/documents/policies/ethical-use-of-student-data/files/22/ethical-use-of-student-data-policy.pdf>. [Accessed: Oct-2019].
- [37] A. Pardo and G. Siemens, "Ethical and privacy principles for learning analytics," *British Journal of Educational Technology*, vol. 45, no. 3, pp. 438–450, 2014.
- [38] S. Slade and P. Prinsloo, "Learning analytics: Ethical issues and dilemmas", *American Behavioral Scientist*, vol. 57, no. 10, pp. 1510–1529, 2013.
- [39] B. Chen and H. Zhu, "Towards Value-Sensitive Learning Analytics Design," *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 2019.
- [40] P. Diaz, M. Jackson, and R. Motz, "Learning Analytics y protección de datos personales. Recomendaciones.," *Anais dos Workshops do IV Congresso Brasileiro de Informática na Educação (CBIE 2015)*, 2015.
- [41] M. E. Gursoy, A. Inan, M. E. Nergiz, and Y. Saygin, "Privacy-Preserving Learning Analytics: Challenges and Techniques," *IEEE Transactions on Learning Technologies*, vol. 10, no. 1, pp. 68–81, 2017.
- [42] D. Lizcano, J. A. Lara, B. White, and S. Aljawarneh, "Blockchain-based approach to create a model of trust in open and ubiquitous higher education," *Journal of Computing in Higher Education*, vol. 32, no. 1, pp. 109–134, 2019.
- [43] A. Mikroyannidis, A. Third, and J. Domingue, 'Decentralising online education using blockchain technology', 2019.
- [44] P. Ocheja, B. Flanagan, H. Ueda, and H. Ogata, "Managing lifelong learning records through blockchain," *Research and Practice in Technology Enhanced Learning*, vol. 14, no. 1, 2019.
- [45] S. S. Oyelere, I. F. Silveira, V. F. Martins, M. A. Eliseo, Ö. Y. Akyar, V. Costas Jauregui, B. Caussin, R. Motz, J. Suhonen, and Ł. Tomczyk, "Digital Storytelling and Blockchain as Pedagogy and Technology to Support the Development of an Inclusive Smart Learning Ecosystem," *Trends and Innovations in Information Systems and Technologies*, pp. 397–408, 2020.
- [46] D. Amo, D. Fonseca, M. Alier, F. J. García-Peñalvo, and M. J. Casañ, "Personal Data Broker Instead of Blockchain for Students' Data Privacy Assurance," *Advances in Intelligent Systems and Computing*, pp. 371–380, 2019.
- [47] M. A. Forment, D. A. Filvà, F. J. García-Peñalvo, D. F. Escudero, and M. J. Casañ, 'Learning analytics' privacy on the blockchain', in *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality*, pp. 294–298, 2018.
- [48] O. A. Naumova, I. A. Svetkina, and D. V. Naumov, "The Main Limitations of Applying Blockchain Technology in the Field of Education," 2019 International Science and Technology Conference "EastConf", 2019.
- [49] S. Khatal, J. Rane, D. Patel, P. Patel, and Y. Busnel, "FileShare: A Blockchain and IPFS Framework for Secure File Sharing and Data Provenance", in *Advances in Machine Learning and Computational Intelligence*, Springer, pp. 825–833, 2021.
- [50] E. Nyalety, R. M. Parizi, Q. Zhang, and K.-K. R. Choo, 'BlockIPFS-blockchain-enabled interplanetary file system for forensic and trusted data traceability', in 2019 IEEE International Conference on Blockchain (Blockchain), pp. 18–25, 2019.
- [51] D. Le Métayer, "Privacy by design: a matter of choice", in *Data protection in a profiled world*, Springer, pp. 323–334, 2010.
- [52] T. Antignac and D. Le Métayer, "Privacy by design: From technologies to architectures", in *Annual privacy forum*, pp. 1–17, 2014.
- [53] S. Slade and A. Tait, 'Global guidelines: Ethics in learning analytics', 2019.