MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
ABDELHAMID IBN BADIS UNIVERSITY - MOSTAGANEM
Faculty of Exact Sciences and Computer Science
Department of Mathematics and Computer Science
Sector : Computer Science

END OF STUDIES THESIS

For obtaining the Master's Degree in Computer Science

Option : **Information Systems Engineering**

Presented By :

Abdallah Lotfi

Selab Linda Hind

THEME :

Design and realization of a real time monitoring system of a student through a Web site.

Supported on : 04/07/2022

In front of the jury composed of :

Dr. Hocine Nadia	MCB	Mostaganem University	President
Dr. Bensalloua Charef	MCB	Mostaganem University	Examinator
Dr. Henni Karim Abdelkader	MCB	Mostaganem University	Supervisor

Acknowledgements

Abdallah Lotfi's Acknowledgements

I would like to start by thanking God the Almighty for his countless blessings, and the opportunity to work on this project and write this thesis.

I would also like to thank my parents for their unconditional support, without you and your constant reminder to focus, I wouldn't have found my passion in this field.

A big thanks also goes to the open-source community, as accomplishing this project without your work is impossible in such a short period of time.

And finally, I would like to express my gratitude to all friends and family who always believed in me.

Selab Linda Hind's Acknowledgements

First of all, we glorify God, the Almighty, for having given us life and strength for all these years and for helping us to realize this project.

To my parents and sisters, who gave me unconditional love, gratitude and supported me throughout my journey.

To my family and friends, who keep encouraging me and believing in me no matter what.

Big thanks to the members of the jury that have accepted and took the time to evaluate our work, and to our supervisor that poured so much time and effort in guiding us.

Finally, I would like to express my gratitude towards those who helped and believed in me since day one.

ملخص

تتكون مذكرة التخرج هذه من جزأين، الأول عبارة عن بحث حول تطبيقات لوحات تحليلات التعليم العالي لمراقبة تعلم الطلاب، لقد بحثنا في المكتبات الأكثر شهرة في هذا المجال لدراسة العديد من البحوث وإيجاد مختلف التقنيات المستعملة في إنشاء لوحات التحليلات التنبؤية والتفسيرية التي تستعمل مختلف البيانات (مثل تفاعل الطلاب مع الدروس) لفهم وتحسين جودة التعليم العالي. من خلال توفير لوحات تحليلات التعليم المخصصة للتعليم العالي، يهدف خبراء المجال إلى مساعدة طاقم التدريس على فهم مختلف نماذج البيانات عبر واجهات سهلة الاستخدام ومخططات بيانية تفاعلية. في الجزء الثاني، سنقوم ببناء لوحة تحليلات في الوقت الآني الخاصة بنا باستخدام بنية معتمدة على بيانات الدروس تم اختبارها جيدا لتحليل بيانات نشاط الطلاب التي جمعناها، ثم نقوم بإيصال البيانات بالإضافة إلى التحاليل في الوقتي الآني باستخدام WebSockets API.

الكلمات الدالة: أنظمة تحليلات التعلم، لوحات تحليلات التعلم، تحليلات التعلم، مكونات المعرفة، الوقت الآني، واب صوكات.

Abstract

This graduation thesis is composed of two parts, first is bibliographic research on the application of Learning Analytics Dashboards (LAD) in Higher Education to monitor students' learning, we searched through the most popular libraries to study multiple papers and find different LA techniques and trends in building predictive or explanatory models that use data (such as students' course engagement) to understand and improve higher education. By providing personalized LADs, LA experts seek to help the teaching staff understand the output of these models through accessible interfaces and interactive visualization. In the second part, we build our own real-time LAD using a well-tested data-driven course structure to analyze the student activity data we collected, and then deliver the data plus the analysis in real-time using the WebSockets API.

Keywords: Learning Management System, Learning Analytics, Learning Analytics Dashboard, Knowledge Components, Real-time, Websockets.

Résumé

Cette mémoire de fin d'études est composée de deux parties, la première est une recherche bibliographique sur l'application des Tableaux de Bords Analytiques d'Apprentissages dans l'´Education supérieure afin de surveiller l'apprentissage des étudiants, nous avons effectué des recherches dans les librairies les plus populaires pour étudier une multitude de papiers et trouver différentes techniques d'Analytiques d'Apprentissages et les tendances de la conception des modèles prédictifs et explicatifs qui utilisent des données (tel que l'engagement des ´étudiants dans les cours) afin de comprendre et améliorer l'´éducation supérieure. En fournissant des Tableaux de Bords Analytiques d'Apprentissages personnalisés, les experts du domaine cherchent `à aider le staff ´éducatif pour comprendre la sortie de ces modèles en passant par des interfaces accessibles et des visualisation interactives. Dans la deuxième partie, on construira notre Tableau de bord en temps réel en utilisant une structure de cours pilotée par les données bien testée pour analyser les données d'activité qu'on a collecté, puis livrer les données et l'analyse en temps réel en utilisant l'API WebSockets.

Mots clés : Systèmes Analytiques d'Apprentissages, Tableaux de Bords Analytiques d'Apprentissages, Analytiques d'Apprentissages, Composants de connaissances, temps réel, WebSockets.

List of Figures

List of Tables

List of Abreviations

CI/CD : Continuous Integration and Continuous Delivery

DBMS : database management system

DL : Deep Learning.

EDM : Educational Data Mining.

HCLA : Human Centered Learning Analytics

HE : Higher Education.

ITS : Intelligent tutoring Systems.

LA : Learning Analytics.

LAD : Learning Analytics Dashboard.

LMS : Learning Management System.

ML : Machine Learning.

MOOC : Massive Open Online Courses.

MVC : Model-View-Controller Pattern

OCL : Object Constraint Language

SWE : Software Engineering

TCP : Transmission Control Protocol

UX : User Experience

Contents

General introduction

For a long time, people have been researching and designing ways to improve teaching and learning, Learning Analytics (LA) and Educational Data Mining (EDM) build on their well-established research by applying data science and taking advantage of the rapidly growing amount of educational data to improve the effectiveness of higher education's learning management platforms by producing personalized insights into students' learning progress.

Although LA can be used with the purpose of supporting both Learners and Educators in a wide range of areas, such as building an intelligent curriculum with adaptive content or even as an alternative to student end-of-course assessment. The focus of this graduation thesis is to solve the problematic of helping a teacher monitor his student's understanding of course concepts and his engagement with course materials.

When developing course materials and exams, teachers usually have a hard time assessing what course materials their students are struggling in and if they could use some help and instruction, they need a Learning Analytics Dashboard (LAD) that helps them in this regard. In this thesis, we will tackle this issue, going over how and which student's data to collect, how to report it to the teacher, and how to allow him to act on these reports.

First, we will start with a scientific research on the subject of monitoring a student's education in using Learning Analytics. This research will contain two (2) chapters :

- The first chapter will look into how the recent rapid advances in Information Technology together with Data Mining and Education Software, has led to the emergence of a new research discipline, called Learning Analytics, with the goal of making Learning Management Systems (LMS) more responsive to the engagement of students.

- The second chapter of our research will give an non-opinionated rapid review of the state of the art of LA in Higher Education (HE), based on the recent research published by publicly recognized LA and EDM experts.

In the previous section, we researched and analyzed multiple Learning Analytics Dashboard. For the last two chapters, we will draw on this research to propose and build our own dashboard:

- The third chapter discusses the problematic that we faced in the design of a LA dashboard in detail, then goes over our proposed solution, from collecting students' online activity to storing and transferring it to the teacher's dashboard, and analyzing this activity.

- The fourth chapter is about the implementation of all that is discussed in the third chapter, we will go into details into how we implemented the logging system, the real-time communication and the Knowledge Components Service. After that, we will show some screenshots of our dashboard, and finally go over the process of deploying it to production.

Chapter I

Learning Analytics

I.1 Introduction

E-learning is the use of hardware, software, and academic content in addition to multimedia and Internet technologies to facilitate the delivery of teaching materials to students. E-learning has a set of features that can improve a learning process's efficacy and efficiency. For example, monitoring students' learning activities using Learning Analytics (LA).

Nowadays, Learning Analytics supports the difficult process of monitoring and evaluating students' learning activities via log data coming from Learning Management Systems (LMS); and visualizations of online activities, which university teachers did not have ten years ago. LA leverages the data coming from the LMS by analyzing it and helping teachers take suitable action to improve learners' experience.

In this chapter, we will go through LA's early history, defining the field and talking about its steps according to some of the latest research; then we go over the most important metrics used in Learning Analytics Dashboards and list the theoretical concepts used in the field.

I.2 What is Learning Analytics?

I.2.1 Definition 1

T. Elias in [1], defines LA as: "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 (EDM), and action analytics."

I.2.2 Definition 2

According to M . Khalil, et al. in their LA review [2], Learning Analytics is a process for collecting, analyzing, and reporting data about students and their environment. As a result, teachers can use LA Dashboards (LAD) to better understand and optimize the learning experience and the environments in which it occurs. This process takes five well-defined steps that are explained in I.4.

I.3 The history of Learning Analytics

I.3.1 The first occurrences of a Learning Analytics concept

Although learning analytics is relatively a new field of research, the concept of learning analytics is not (Figure 2). Before the emergence of learning analytics, the use of various types of learning technologies and learning data has been already known in educational research. Much of the data mining tools and approaches that are now widely employed in learning analytics research, for example, social network analysis or discourse analysis.

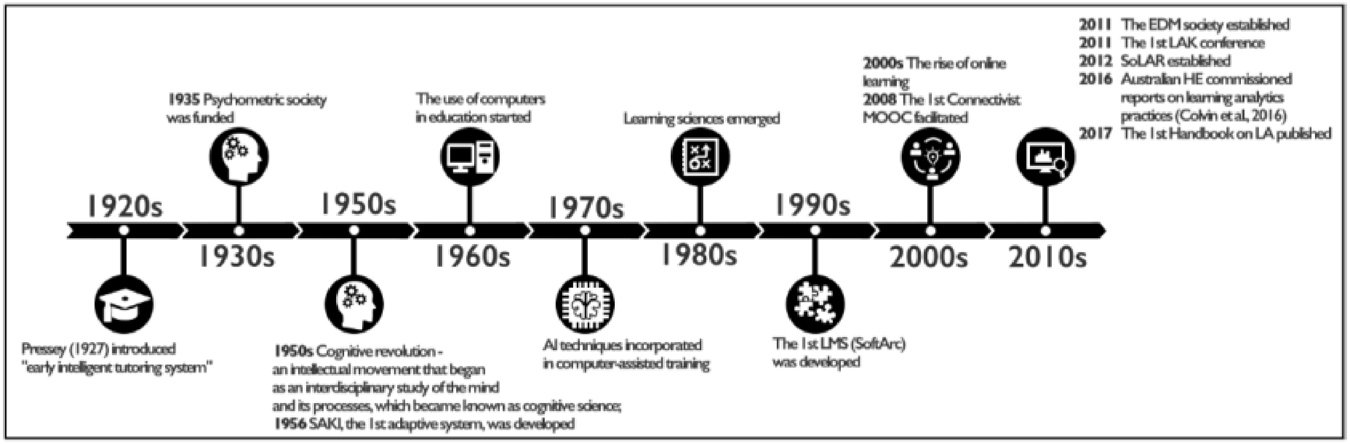


Figure I.1: The genealogy of Learning Analytics[2]

.

The first automated teaching machine was developed in the 1920s by Pressey. His work can be argued as the start of intelligent tutoring systems (ITS), one of the key areas upon which learning analytics draws.

Furthermore, cognitive science originated in the work of Miller in the 1950s and new advances in computer science and artificial intelligence was another critical influence.

In 1956, the first adaptive teaching system SAKI (Self-Adaptive Keyboard Instructor), was developed for teaching keyboard skills by Pask, McKinnon-Wood, & Pask in 1961. SAKI improved learning rates by matching task difficulty to learner performance. This demonstrated an important finding of how student learning can be supported through the use of technologies at scale which had profoundly shaped the development of modern educational technology and, thus learning analytics, is the growing realization of the benefits of personalized instruction.

The growth of online and distance education further contributed to the development of learning analytics [3]. Starting with the use of the postal services in the late 19th Century then various technologies including radio, television, video, CD, DVD, and now more commonly, the Internet[4].

The development of World-Wide Web in the 1990s gave birth to Web-based distance learning systems which in turn evolved into modern-day online learning [3].All this led to the creation of Massive Open Online Courses (MOOCs), a particular form of online learning in which thousands, and even more, students engage in distributed, online learning.

The expansion of the Internet during the 1990s and 2000s led to web-based distance learning technologies known as Learning Management Systems (LMS).

I.3.2 The emergence of LA

The appearance of academic ranking of world universities in the early 21st century, mainly Shanghai Jiao Tong Ranking in 2003 and Times World University Rankings in 2010, helped a lot of universities attract students and increased the academic competition between them to a new level, the researchers and administrators at universities used these rankings as a benchmark to help increase the quality of education and staff, and the yearly research output.

This competition and the increasing number of educational datasets showed researchers the unused opportunity in the field of learning theory and educational data mining technology, which birthed the field of Learning Analytics. This increased interest in LA revealed itself with the number of dedicated conferences, workshops and white papers. Examples mentioned in the paper are:

- The 1st Learning Analytics and Knowledge conference in Banff, Canada, 2011; the goal of this conference was to establish this emerging research field and to reveal the possibilities it provides in understanding students' learning through the use of machine learning, data mining and data visualization methods.
- The 4th International Conference on Educational Data Mining 2011 in Eindhoven, Netherlands; which focused on the topic of mining students' data, whether it is assessment data or browsing activity, it brought together researchers from computer science, education, psychology, psychometrics, and statistics backgrounds to analyze large datasets and answer educational research questions.
- The 1st DATATEL workshop on Educational Datasets for Technology-Enhanced Learning at in France.

I.4 The Steps of a Learning Analytics Process

Figure I.2: The five steps of the analysis process.[2]

.

In our study into Learning Analytics in Higher Education, there is one main actor that the data is based on, the student who uses the Learning Management System to learn. The objective of LA is to improve his learning experience and provide helpful feedback to teachers and university staff on his performance and achievements.

The data produced by students passes through the lifecycle of the LA process [2] (Figure 1). First comes capturing, then reporting, prediction, acting, and finally refining.

I.4.1 *Capturing*

LA requires a wide range of data for the analysis and predictions to be accurate, if we take grades as an example, the final exam grade alone is not sufficient to provide helpful insight inside the course. Most university courses choose to divide their course grades into projects, attendance, tests, homework and exams, etc. the more varied grades are, the more accurate the generated models are. The mining of real-time data comes with its own issues, however; such as storage and granularity.

I.4.2 *Reporting*

Once data is stored, it is time to use analysis tools to examine the information and identify patterns. Nowadays, LA dashboards (LAD) are widely used to showcase what was learned through analyzing data, which makes reports easier to read and offers better understanding of the students' progress. In their study on LAD and reports, *R. Paiva, et al.* [5] found that the use of visualization tools is more effective at making teachers understand the LA reports.

I.4.3 *Prediction*

The development of predictive models is common for LA Dashboards, as they rely on multiple models for each time period (quizzes period, essay period, exams period, etc.) and each set of data such as students' engagement with the course, or quizzes' weekly results. For example, regression could be used against class attendance data (or the average time per online LMS session in case of e-learning) to see the likelihood of a student returning the following year. An important variable in using prediction models is *frequency*, in the example above, the model might only be needed once a year, but if we are trying to predict the likelihood of students' failure, then the model should be run at least weekly or monthly against each new provided dataset.

I.4.4 *Acting*

This step is the end goal of every institution, taking action and making interventions based on the predictions and reports, Whether the action is sending warnings to students who are at risk of failing, or changing course material or a teacher in case a high percentage of students are failing. In a real-time LA reporting system, actions and interventions can even be taken during a teaching session.

According to Susanne Narciss and Katja Huth in their research on designing an informative tutoring feedback framework [6], real-time feedback to both teachers and students is another form of intervention that is well-established in the E-Learning field. The feedback could take the form of hints on how to correct students learning process, possible sources of information that the teacher could provide in case students are stuck on a problem, etc.

I.4.5 *Refining*

Every analytics project needs to update regularly, whether to add new parameters or improve the process to avoid overfitting and underfitting and reduce standard errors in the predictive model.

Another form of refining the process is by determining if the actions taken in the previous step were helpful or not. It is possible to determine if a teacher's intervention was useful to the students by looking at the end-of-semester results of students, however that is the old approach. The appearance of LMS and Intelligent tutoring Systems (ITS) has made it possible for universities to analyze the log data of students through educational data mining to understand how effective each intervention was.

I.5 Metrics Used in Learning Analytics

Building a LA tool means the collection of a large amount of data, knowing what data is most important and what should be included in the LA dashboard is challenging. Despite the fact that each course is different, there are several metrics that should be included in every dashboard.

- Logins
- Total time spent on activities
- Average time per session
- Test and Exam Scores
- Most Viewed Course Materials
- Time Spent in Every Course
- Number of Passing and Failing Students
- Behavioral Patterns in Students' Learning Experience
- Outcome Predictions
- Personalized Course Metrics

I.6 Theoretical Notions

I.6.1 Machine learning

An application of artificial intelligence (AI) that can give the ability for a machine to learn and improve automatically through experience and by the use of data, in order to make predictions or decisions without being explicitly programmed to do so. Machine Learning is applied when classic algorithmic and software engineering reaches its limits. Its core concept consists of observing data thoroughly, be it through experience, or instruction. It mainly looks at recurring patterns of a given data and extracts it to detect it easily on every data given. [1]

I.6.2 Data mining

An Interdisciplinary sub-field of computer science and statistics that involves extracting, discovering and deciphering data patterns through computation using tactics that combine technology concepts such as artificial intelligence, database systems, machine learning, and statistics. In data mining, the overall goal is to extract concrete information (with intelligent methods) from a data set and transform the information into a comprehensible structure for further use. Factors that come into play in the process

[1]https://www.dominodatalab.com/data-science-dictionary/machine-learning

of data mining include managing databases and data, pre-processing the data in question, referring to helpful constructs such as models and inferences, and more.[2]

I.7 Conclusion

In this chapter, we presented an overview of LA and defined its process, talking briefly about its history, the LA process is highly efficient in meeting the needs of students and teachers. Moving forward, it has the potential to fully maximize students' success if the right tools and steps are used to realize.

LA is garnering increasing popularity in data and machine learning, as well as educational science, and is thus becoming an important element of modern computer-based education systems. In the next chapter, we will talk extensively about the most recent research into LA and EDM, studying what modern experts have achieved in the technology.

[2]https://blog.amt.in/index.php/2020/02/27/introduction-to-data-mining/

Chapter II

State of the Art

II.1 Introduction

The growing demand and usage of online education platforms for higher education over the last decade has decreased the face-to-face interactions between students and teachers, which made teachers' job in monitoring and judging students' advancement and understanding of course materials harder. This led to the increased need for Learning Analytics (LA) tools to study the massive amount of data collected about students' activity, and ML models with the goal of predicting student performance.

In this day and age, the use of both predictive and explanatory data analytics algorithms in improving decision making is already widespread in many sectors. In this chapter we will discuss recent research studies into learning data analytics, and examine the state of the art in the field of LA, looking into how researchers used data-driven solutions to improve the learning process in Higher Education, by determining how students perceive online courses and which activities have a substantial effect on their success.

II.2 LA and EDM Techniques mentioned in papers

The most recent review of EDM applications in Higher Education was done in 2019 by Aldowaha, et al. [8], in their research into 402 articles, they found that *Classification* is the most commonly applied data mining techniques in Higher Education LA; by a percentage of 26.25%. It is used to predict students' performance, potential dropout and even achievements and final grades.

The next most used technique is Clustering, by a percentage of 21.25%. In Higher Education, it is used for the identification of patterns in students' activities and grouping said students based their individual learning characteristics to help the teaching staff allocate them to the most suitable courses and help potential dropout.

Other Data Mining techniques that are used in Education are visual data mining (15%), statistics (14.25%) and regression (10.25%).

II.2.1 Automatic Predictive Analytics via Dashboards

Automatic Predictive Analytics is a technique that provides teachers with the ability to use the power of LA by delivering real-time feedback about their students' performance. Predictive Analytics considers not only grades, but also prior academic history, and students' effort as measured by LMS interaction, plus student's demographic data such as age and gender. The results are delivered to teachers via a personalized dashboard that shows how each student is doing. As an example, Course Signals [8] which is considered the first LA dashboard; used predictive models to label students' risk status using traffic light signals.

Figure II.1: Course Signals at Purdue University showing green (safe) and yellow lights (borderline safe)
.

Another Example is E2Coach, which is an adaptive advice system developed by the University of Michigan that provides a smart dashboard tailored to each student with recommendations and guidance on learning resources to consider taking, it also provides personal learning goals, messages alerting the student of bad performance and even supportive messages. The dashboard also provides helpful learning strategies that help students perform better and complete their courses [10].

OU "Open University" Analyse [1] is another example, the tool's main goal is the early prediction of under-performing students where an intervention can have a serious impact on their learning. It uses four predictive models, a Bayesian classifier, k-NN, classification and a regression tree to analyze students' demographic data; and the k-NN algorithm with the data coming from the LMS. This method reported 90% accuracy at predicting at-risk students by the end of the semester [11].

[1]`https://analyse.kmi.open.ac.uk`

Figure II.2: OU Analyse: ML powered system to identify students at risk of failing

.

Several commercial products such as D2L[2], Blackboard[3] and Lambda Solutions[4] also provide Predictive Analytics consulting and products for HE Learning management systems [12]. that include custom reports with embedded visualizations directly in the LMS, which shows how well established this research field already is in practice.

The goal in predictive modeling is to create models that predict values or classes of new students' data, based on already verified course observations that could be obtained from previous semesters of the same course, or historical log files of each student (which we refer to as training data). According to C. Brooks and C. Thompson in the fifth chapter of The Handbook of Learning Analytics, titled "Predictive Modelling in Teaching and Learning" [12], when building a model, a developer should collect as much data as possible, as feature selection is possible. The possibility of creating a model that only relies on a single attribute that correlates with the outcome is improbable.

Although missing data can be dealt with by removing rows or columns of the dataset, in Learning Analytics, data is limited by what the LMS log files provide, so removing data may negatively impact our models. The second solution is to replace the missing value with the mean of the values from other rows (instances of data), or if possible, fill it with data copied from similar rows.

[2]https://www.d2l.com/
[3]https://www.blackboard.com/
[4]https://www.lambdasolutions.net/

A general assumption about predictive models is that closely correlated attributes to a target variable means we can easily make the prediction and that it always holds true, however this assumption does not always apply, for instance, a dataset that shows students who score high in their core (first year) curriculum' calculus and physics class are always completing their masters' or engineering five-year program, which may not work in practice.

Several well tested algorithms are used to train predictive models such as *Simple* and *Multiple Linear Regression* for continuous numerical outputs, and several classification algorithms like *Logistic Regression*, *Nearest Neighbors* and *Decision Trees Classifiers*. Each of these algorithms can be implemented into an LA dashboard with tunable parameters.

After training the model comes the evaluation phase, a common practice predictive analysis is to use *k-fold cross validation* in order to obtain a large testing set to evaluate the quality of the model using measurements such a prediction accuracy, precision and recall. In the educational field, it is often good practice to used data of students from previous semesters as a training.

II.2.2 Explanatory Data Analysis

According to Rain Liu and Kenneth R. Koedinger in their article [13], The vast majority of the LA field has focused on researching and building predictive models, but they argue that Explanatory models could benefit from more focus. In their article, they discussed examples of educational data mining that let to improvements to students' learning outcome.

One of the methods discussed in the article is the use of *Data-Driven Cognitive Model Discovery* to identify problematic elements in certain tasks given to students (quiz questions, exam exercises, etc.) with the use of *Knowledge Components Modeling (KC Modeling)*, if a difference is found between how students responded to two closely related tasks, the teaching staff is notified that the harder task requires knowledge that may not be present in the easier one.

This is achieved using KC Modeling, an advanced cognitive model accomplished by splitting a certain exercise or problems into multiple steps, each step requiring certain knowledge that will help students to solve it, each student is given several chances (opportunities) to demonstrate whether they learned a certain knowledge. After that, a KC model that maps each step to one or more knowledge components is built to help teachers visualize (using a custom LA dashboard) the areas where their students are lacking by calculating the error rate for each attempt a student has taken to solve a step of a problem, which is shown in the figure below (taken from a DataShop[5] cheat sheet [13]).

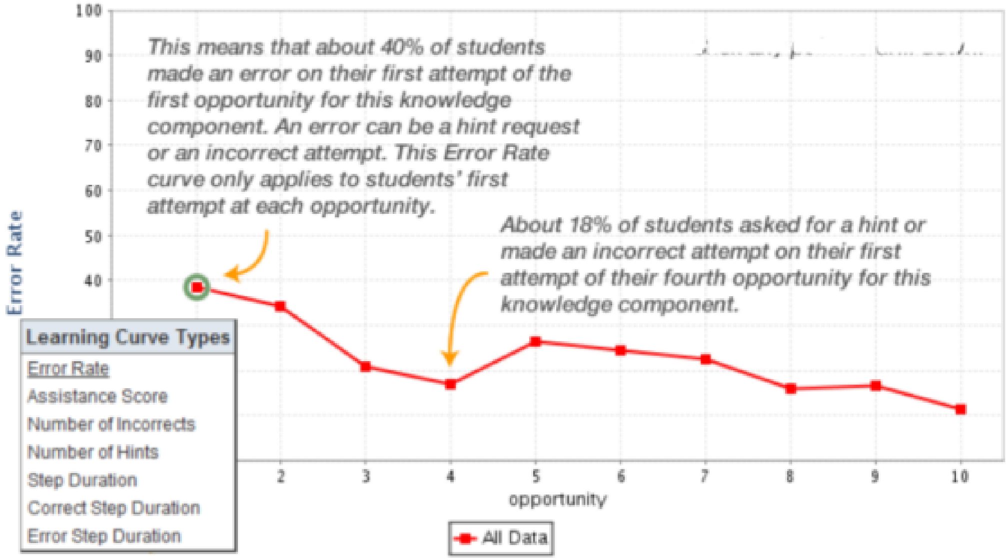


Figure II.3: A visualization of students' learning curve accomplished using KC Modeling

While visual explanatory analysis seems like the way to go, researchers M. A. Chatti, et al. argue that human centered LA (HCLA) matched with the Visual Analytics (VA) approach that emphasizes the human factors in LA [15], is the solution to the lack of acceptance and adoption of LA in general, and Explanatory Analysis specifically.

Although visualization in LADs is a popular approach for representing students' performance data, these dashboards are static in general and don't offer the teaching staff much interaction, which is hard to implement in practice, especially in a personalized dashboard that carters to the needs of different courses and learning approaches.

[5] 5https://pslcdatashop.web.cmu.edu/about/

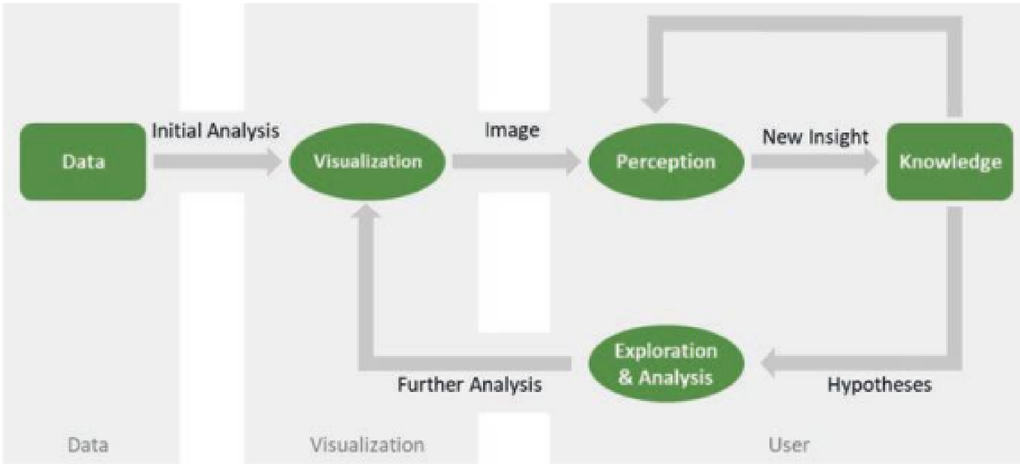


Figure II.4: The Loop of Visual Analytics[16])

The concept of HCLA is displayed in the figure above, where LA meets the needs of teachers since they are involved in the process where visual analysis is the first step. Through exploration of the initial VA, the teacher can gain new insight and use it to further improve the process, the system then enters a loop where it uses both previous iterations and the teacher's hypotheses to derive additional analysis.

II.3 Learning Analytics in Higher Education

In their book about the adoption of LA in Higher Education (HE); David Gibson et al. [17], talked in details about the five major HE domains that could be improved using LA and outlined fifteen tactical methods (three for each domain) that could be used to improve each domain, these domains are the following: (1) acquiring students, (2) promoting learning, (3) offering timely relevant content, (4) using up-to-date delivery methods and finally (5) supporting alumni networks.

In our research into following the education of a student in real-time through one study subject over one semester, we will only focus on the second domain, Promoting Learning, where the authors give us an idea on the techniques and tactical methods used to improve students' learning experience, the techniques mentioned are:

II.3.1 Adaptive Support

In this method, Learning Management Systems use students' interaction and learning history to learn the preferences of students, and either automatically or manually adapt, by providing self-assessments, delivering adaptive recommendations, and producing personalized analyses of their learning activities. according to the book, this is a key mechanism and one of the primary benefits of using LA in Higher Education.

II.3.2 Proactive Retention Management

As several researchers have pointed out, students with high attrition risk can be identified early and receive targeted preventive interventions.

Proactive retention is an eminent subject in the literature since it balances benefits to both learners and the educational system. It underlines for learners that analytics may assist decisions and develop human potential while also supporting organizational and educational efficiencies by saving time and money.

II.3.3 Personalized Communication

Learning materials can be targeted to communicate with students based on learning characteristics, level of achievement, and aspirations for attainment with suitable adoption of learning analytics.[18]

This includes using network analysis to understand social and cognitive relationships in learning and creating conversational agents (Artificial Intelligent AI agents) to enhance learning and the educational journey through higher education [19]. Taken together, these and other technological advancements have facilitated personalized communication which is essential for student engagement and retention. In fact, this leads to improve the student experience (and retention), teachers' relations and students.

II.4 LA challenges and opportunities

Despite all of the robust tools made available for educational researchers, the learning analytics community is still facing a number of challenges when building, validating, and applying models. We identify different areas that could use investment in order to increase the impact that predictive modelling techniques can have:

- Support non-informaticians in explanatory modeling activities. The field of learning analytics is highly interdisciplinary, and educational researchers, psychometricians, cognitive and social psychologists, and policy experts typically have strong experience with explanatory modeling. Providing support in the application of predictive modeling techniques, whether through the innovation of user-friendly tools or the development of educational resources on predictive modeling, could further diversify the set of educational researchers who use these techniques.

- Create community-led educational data science challenge initiatives: it is typical for researchers to work on the general theme but using different datasets, implementations, and outcomes, resulting in difficult-to-compare results. Moving toward a common set of outcomes, open and more detailed data, and shared implementations to compare the efficacy of techniques and the applicability of modeling methods for given problems might be beneficial to the community. Furthermore, this approach has already been valuable in similar research fields and the broader data science community.

- Engage in second order predictive modeling : in LA, second order predictive models include historical knowledge about the effects and intervention in the model itself. As an example, a first order predictive modeling would use student interactions with content to determine drop out. While a second-order predictive model would include historical data on the effect of an intervention (such as an email invitation).

It is essential to turn to modeling the effectiveness of interventions when multiple interventions are available and personalized learning pathways are desired.

Although the multidisciplinary nature of the learning analytics and educational data mining communities, there is still a significant need for bridging between the various researchers involved. An interesting underlying theme at learning analytics conferences is discussions about the roles of theory and data as drivers of educational research. Have we reached the point of "the end of theory" [22] in educational research? Although unlikely, this question is relevant to the subfield of predictive modeling of teaching and learning. It is important to discuss more explicitly the goals of research programs in this area in order to better guide methodological choices between explanatory and predictive modeling techniques.

II.5 Conclusion

In this chapter, we discussed the techniques applied in LA and EDM in higher education, and found that various solutions are suited to be implemented in the field, we also studied several implementations and dashboards to see their effect on students and higher education on general. We also learned of the impact LA has on learning process by making a teacher's job of evaluating students' understanding of the course much easier [13].

Chapter III

Design of a real time monitoring system of a student through a Web site

III.1 Introduction

In this chapter, we will present the design steps of our Real-time Learning Analytics Dashboard, we are going to discuss our proposed solution in details. Then, we will describe the dynamic nature of our Learning Analytics Dashboard and how Teachers can interact with it with the help of Unified Modeling Language (UML) Behavioral Diagrams. Finally, we use the UML Class Diagram to form a an architecture of our application.

III.2 Proposed Solution

Following the research onto the state of the art of LADs, we have a general idea into how a Learning Analytics dashboard could be designed.

To facilitate the process of writing our proposition, and making it easy to understand for the reader, we will first define *the purpose of our dashboard*, what will be the its benefit once the desired outcome of this design phase is over? Next, we will go in details into *what base tool and analytics method we chose to monitor student's education*. Finally, in an attempt to defend our proposition, we will *explain why we choose this method*, its pros and cons, and how it will be beneficial to the teacher.

III.2.1 Purpose of our Learning Analytics Dashboard

As the theme of our graduation project suggests, we are trying to design a web dashboard, that allows a university teacher to monitor his student's learning progress in one specific course through out a whole semester. What the research phase told us is that an important part of monitoring a student is quantifying his learning progress and giving it a value. While most universities use grades and credits for this purpose, and for enabling said student to graduate in his studies, our goal for this dashboard is not to help a teacher give his students a passing grade or to see if they qualify for progressing in their education, it is only to help them monitor his students' learning, and their degree of awareness of course modules.

In most universities, using quizzes, assignment, homework, essays and group projects is critical for this purpose, and since they are all well tested methods to quantify student's learning, we chose to pick one of them for our work. But, in choosing one of them, came the critical factor of *"Real-time Monitoring"*. It is obvious that giving a teacher the ability to monitor home-works and group projects that may last for a week or even up to the whole semester is impractical, that is why we narrowed our choice to working either with Quizzes or Assignments.

To maximize the gain the teacher gets from the LAD, and after studying and evaluating the tools and algorithms we could use on top of our two possible choices, we came into the conclusion that both possibilities offer a good base for our solution, so we ended up choosing Multiple Choice Quizzes, as having multiple quizzes through out a semester is a practical method to assess learning progress, and it gives us a variety of options to choose from when it comes to working with student's quiz results and all its associated data. Taking into consideration that a regular educational semester lasts for up to four months, then having 2-3 quizzes every month will give us a good dataset (about 8-12 quiz results for each student in one semester) to aid the teacher in his mission.

The next step in finding a solution for us (in addition to the research that was done in the first phase of this project), was to find a fitting algorithm for our chosen Quiz evaluation method. In doing so, we first had to create an architecture III.4 of the quiz system and describe the behavior III.3.3 of our teachers and students around this system. This important step gave us a good outlook of the structure of our data, making finding a suitable algorithm much easier, which we will explain in the next title.

III.2.2 Knowledge Components

During the research phase, more specifically in the second chapter, where we talked about Explanatory Data Analysis II.2.2, we referred to the Knowledge Components Modeling Technique, that uses Data-Driven Cognitive Model Discovery to accurately draw a *Learning Curve* for each student. A learning curve, as shown in Figure II.3, is a visual representation of a student's error rate trying to solve a certain problem, achieved using KC Modeling.

In our proposed architecture of an LAD featuring a Multiple Choices Quiz System, where a teacher monitors his students in real-time, this modeling technique seems like a perfect fit to help us map every Quiz Question to one, or multiple lessons in a course. To draw a simple example of what this may look like in practice, we made the figure below III.1 that illustrates how a Software Engineering (SWE) course may look like in our system, in addition to its knowledge components.

This map drawn between Quiz Questions and Lessons via Knowledge Components will make it easier for a teacher to judge his student's understanding of course materials. For example, and as shown below, The Object Constraint Language (OCL) lesson's KC, named "data-types", is shared between at least two questions in "Quiz 01", by following this pattern and after multiple quizzes, the teacher will have an better idea on what his students are struggling in, and he can take action based on this new knowledge.

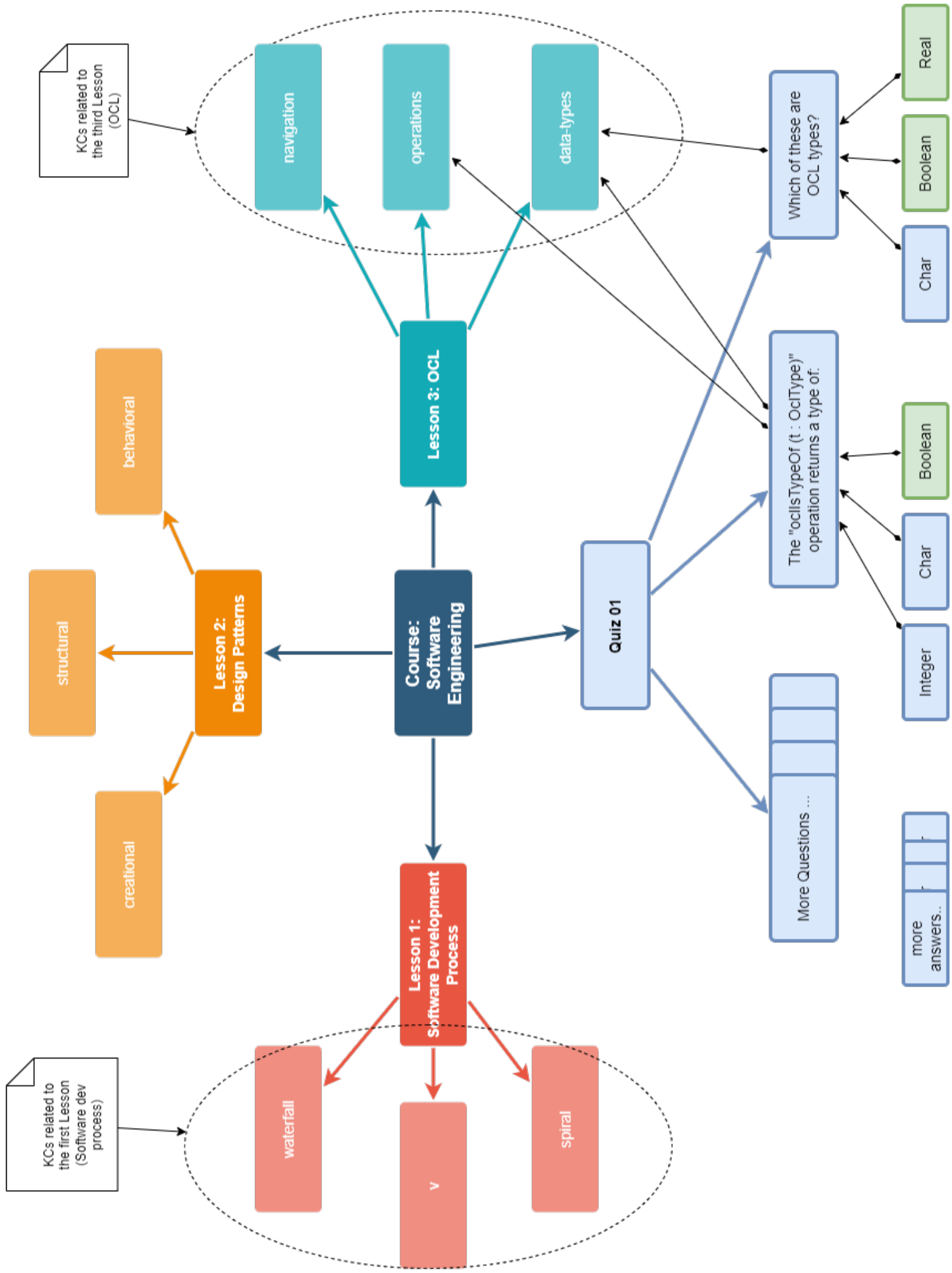


Figure III.1: Illustration of a SWE course with its Lessons, a Quiz and KCs

Using the structure above, we can effectively monitor a student's learning process, we can treat this analysis method as a service, or a feature of our dashboard, and we can decompose this service into components to further organize it:

Student's Knowledge Component awareness in a quiz : In this component, we can see if a student is aware of certain course's knowledge components or not, which will give the teacher an in-depth understanding of his student's learning progress.

Student's Knowledge Component rating in a quiz : It is important to give every student a Knowledge Component rating based on his performance in every quiz, and every quiz question related to this KC, this rating will help us draw a radar representation of a student's KC understanding in a quiz, which is a useful way to display multivariate observations with an arbitrary number of knowledge components, and also makes the teacher's monitoring task easier.

Student's Knowledge Component overall rating : Since every knowledge component belonging to a lesson or a quiz question, also belongs to the related course, this service component, goes over all the course's knowledge components and give a student a rating of his understanding, which is our method for giving the teacher an end of the semester performance review of his students.

Quiz's general Knowledge Component error rates : An error rate is the obvious way of evaluating how students performed on knowledge components, this is achieved by going over every KC in a quiz, fetching every score of a related question, then calculating a general error rate, that will help the teacher in splitting knowledge components in case of an unexplained huge difference in error rates of questions that share the same KCs.

Notify a student if unaware of Knowledge Components : When or after a student passes a quiz, a teacher is able to easily identify students that are struggling in a Knowledge Component, and thanks to this service, he should be able to notify them in real-time.

Fetch possible splittable Knowledge Components in a Course : Ideally, when the semester is over, and with the help of our dashboard, a teacher is able to fetch the the Knowledge Components that are splittable, our dashboard judges a Knowledge Component as "splittable" if it has questions where the error rate of one is significantly higher than the others, indicating that students have found one of the questions much harder to answer. We call this difference the "split percentage", which we will give the teacher the ability to specify in the fetching/splitting process.

Split Course Knowledge Component : After fetching possible splittable Knowledge Components, a teacher (guided by the dashboard in a multi-step process) and thanks to this service component, should be able to follow well defined steps that will help him in splitting a knowledge component, and assign the newly created KCs them into all the lessons and questions of the old KC.

Export All KC scores data : Exporting student KC awareness data into a spreadsheet to give the teacher a portable excel file containing a list of students with their Quiz scores, KC scores.

These service components are not independent, they call on other services, mainly those that are related to quizzes, and calculate quiz scores for students and pass percentages, check quiz statistics on every update to the database by a student, or check if responses are correct, etc.

III.2.3 Advantages

- As shown in the figure III.1, and explained above, mapping Lessons to Quiz Questions is beneficial for making teachers' job easier in assessing how much aware a student is of his lessons.
- Drawing a learning curve and giving the teacher a general error rate for quiz questions helps the teacher to better plan future quizzes by focusing on what students are lacking in.
- The map between questions and lessons via KCs allows a teacher to be sure of his lessons structure and the Knowledge Components added to each lesson and quiz question. In case two questions share the same KC, but have a very noticeable difference in error rate, the teacher can draw the conclusion that maybe his KC structure is inefficient.
- The map also allows us to draw a Radar Chart representation of a student's KC understanding in a quiz, as shown in Figure IV.4.3, which gives value to his awareness of every KC.

III.2.4 Disadvantages

- As mentioned in Section II.2.2, KC Modeling is an excellent tool for monitoring student's learning process and helping teachers develop better course materials. But what it makes up for in explanatory analysis, it lacks in the predictive approach.
- According to The DataShop cheat-sheet [13] and its associated papers, the number of iterations it usually takes to build a perfect KC model differs for one course to another, but it always requires a good amount of testing and mistakes to achieve.
- Splitting Knowledge Components is a critical step in building a good fitting KC Model, but it can not be achieved on just a few responses with one or two questions, a KC must have a large amount of related questions and responses for the teacher to have a good reason for splitting it.

III.3 Requirements Analysis

III.3.1 Actors

Actor	Description
Teacher	A teacher manages his course materials (sections, lessons, quizzes) and the associated Knowledge Components. He can monitor his students' activities to gain a better understanding of their learning.
Student	A student can access the website to access course materials and take quizzes.

Table III.1: Actors

III.3.2 Context Diagram

In the Context Diagram, we will present our system's environment, it is the perfect tool for introducing and managing requirements. As stated in the last title, our system has two actors. All our data will be stored in a relational database. and this data will be communicated to different services using the websockets API, and the student's activity will be constantly logged.

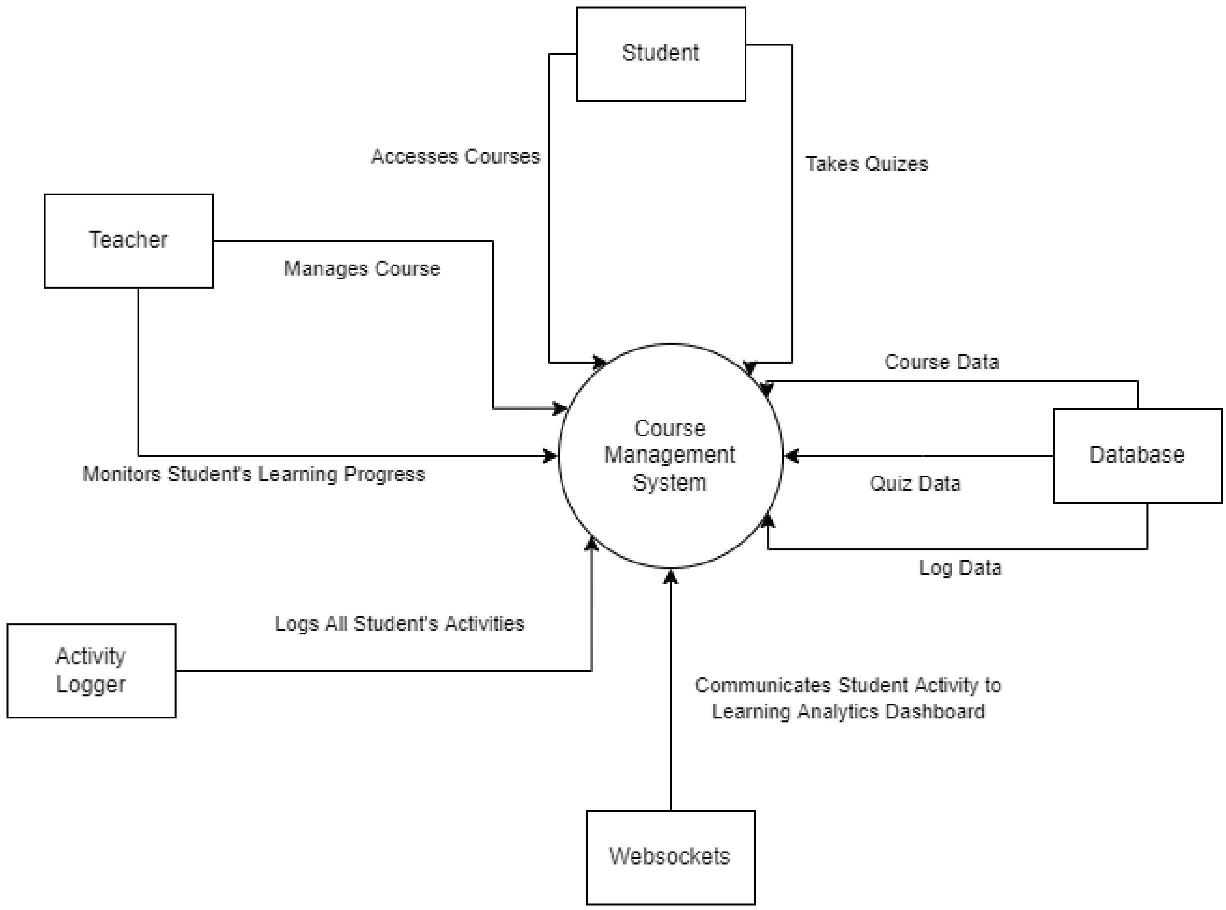


Figure III.2: Context Diagram

III.3.3 Use Cases

Global Use Case

In this diagram, we present a broad picture of how our two actors are expected to behave in the application. It can be seen that the teacher can manage and monitor his course, quizzes, and Knowledge Components. While the student has access to his course materials, and can take his assigned quizzes.

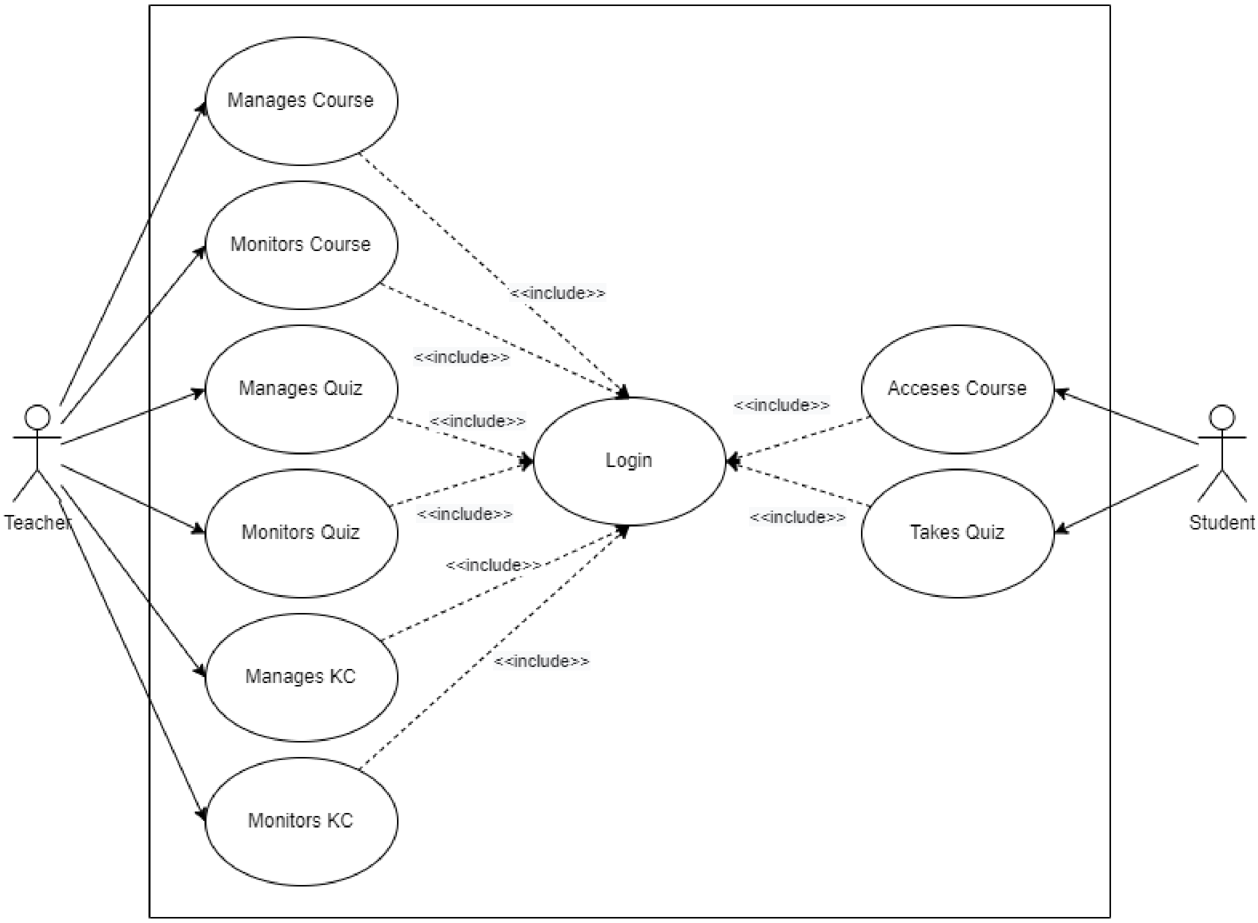


Figure III.3: Global Use Case Diagram

Login Sequence Diagram

In this diagram, we go into how the Login interaction happens in our dashboard, when a user first visits the login page, he is invited to fill his login information, the authentication controller validates the request and either redirects him to the appropriate page (if he is a student or a teacher), or sends a failed login response.

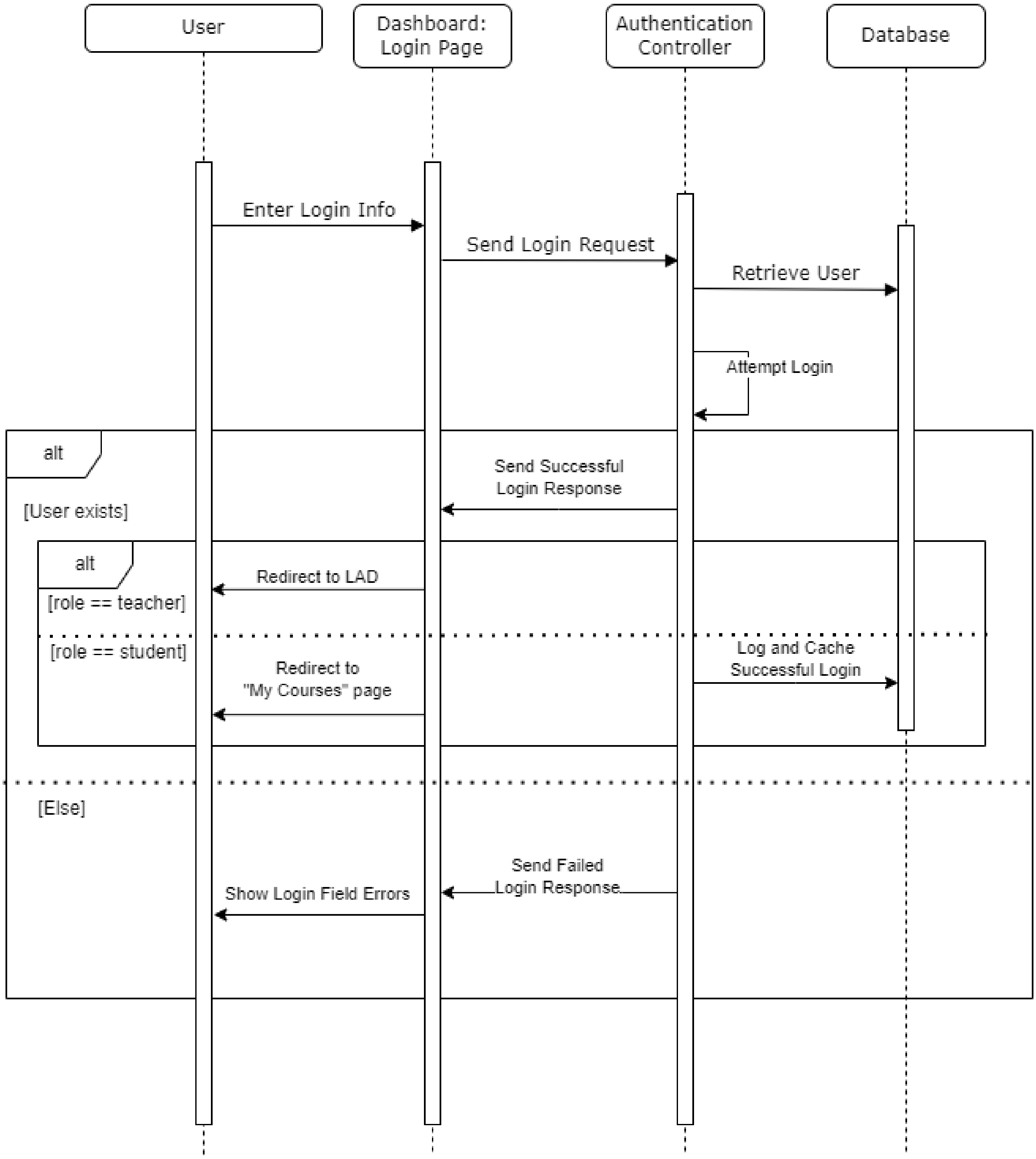


Figure III.4: Sequence: Authentication Process

Course System

Teacher Manages Course

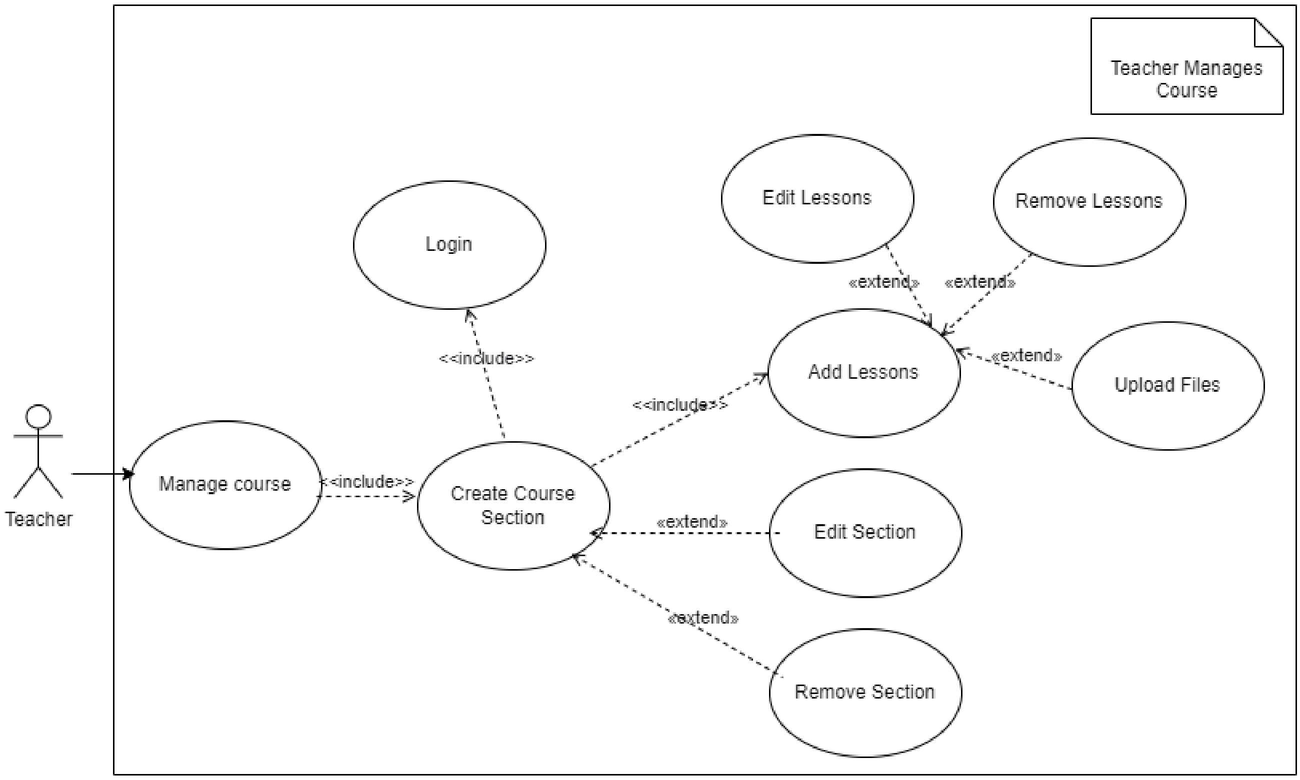


Figure III.5: Use Case: Teacher manages course

Use Case	**Teacher Manages Course**
Description	Teacher manages a course
Actors	1. Teacher
Preconditions	1. Teacher must be logged in 2. Teacher must have the permission to manage courses 3. In case of modifying a section, lessons or KCs, the course must not be active yet.
Main Scenario	1. Teacher accesses the dashboard 2. Teacher clicks Create New Section in Course Navigation Bar 3. Teacher fills necessary information to create a new section 4. He is redirected to Course Page 5. Teacher clicks on the new section just created in the course navigation bar 6. Teacher clicks Add Lesson 7. He fills in the necessary information to create a new lesson 8. Teacher uploads the lesson files. 9. He saves the changes. 10. Teacher notifies students of the newly created lesson.

Alternative Scenario	8. While adding a new lesson, the teacher doesn't find the KC he wanted 6. Teacher is prompted if he wants to create the new KC 7. Teacher wants to create the KC, he is redirected to the "manage Course KC Page" 8. Teacher creates new KC. 9. Teacher is redirected back to Manage CoursePage. 10. Teacher adds the newly created KC to the new lesson. 11. Teacher saves the changes
Post-conditions	1. A validation message is displayed 2. Teacher is redirected to the stats of student's views page.
Exception	1. A field is not valid. 2. Resume to refilling information's step.

Table III.2: Use Case: Teacher Manages Course

Teacher Monitors Course

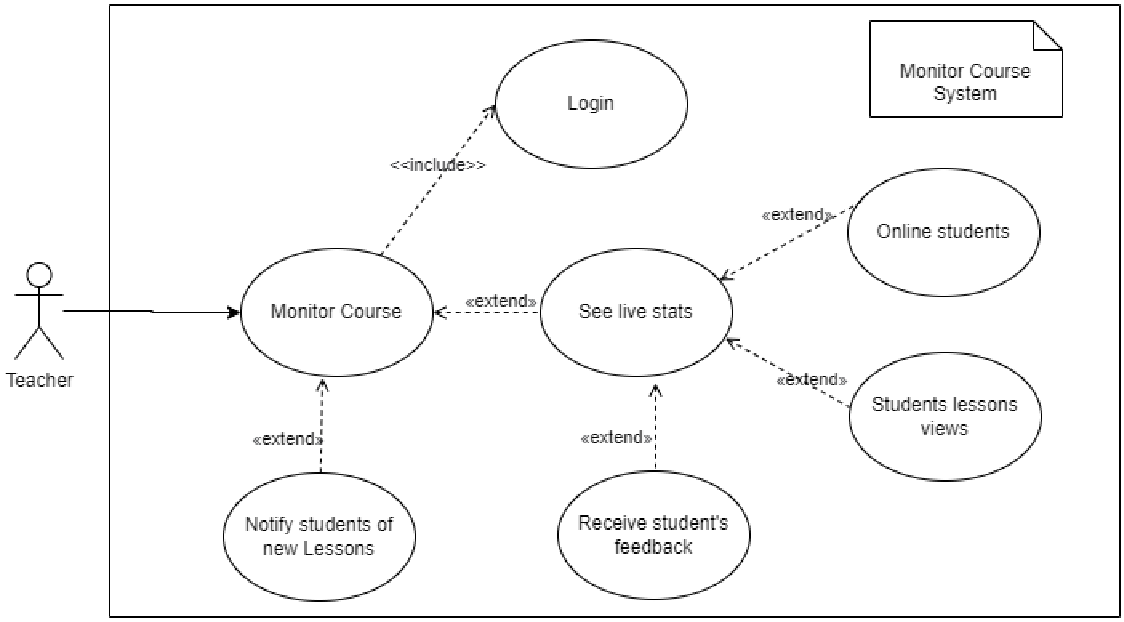


Figure III.6: Use Case: Teacher monitors course

Use Case	**Teacher Monitors course**
Description	Teacher monitors a course
Actors	1. Teacher 2. Student
Preconditions	1. Teacher and Student must be logged in 2. Teacher notified student about the new lesson 3. Each lesson must have its KCs

Main Scenario	1. Student accesses the lesson. 2. Teacher sees students online and those accessing the lesson in real-time 3. Student can download lesson files and add feedback. 4. Teacher can understand student's struggle through their feedback.
Post-conditions	1. Every student who accesses a lesson is recorded

Table III.3: Use Case: Teacher monitors course

Student Accesses Course

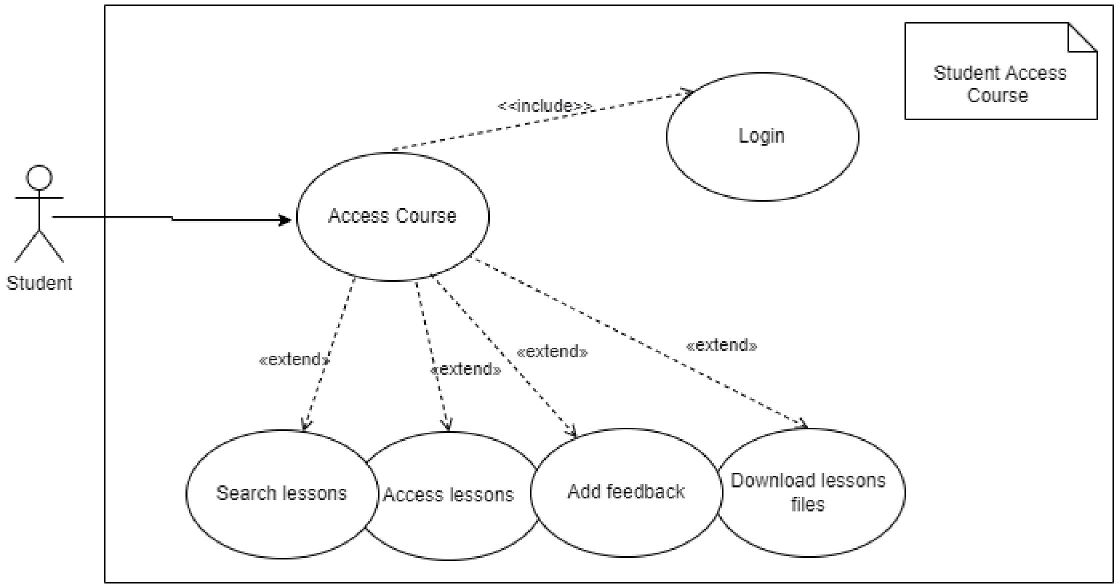


Figure III.7: Use Case: Student accesses course

Use Case	Student Accesses Course
Description	Student accesses a course
Actors	1. Student
Preconditions	1. Student must be logged in 2. Student must have the permission to access courses.

Main Scenario	1. Student accesses the course page 2. He searches from the sections displayed the one that contains the lesson he wants to access 3. Student clicks the target lesson 4. He is redirected to Lesson Page concerned. 5. The lesson is displayed, he can download the lesson file needed if it exists. 6. Student can add his feedback about the lesson 7 Teacher reads the student's feedback, notices the one's struggling and also sees the stats of the student's views on the lesson.
Post-conditions	1. Every student who accesses a lesson is recorded.

Table III.4: Use Case: Student Accesses Course

Quiz System

Teacher Manages Quiz

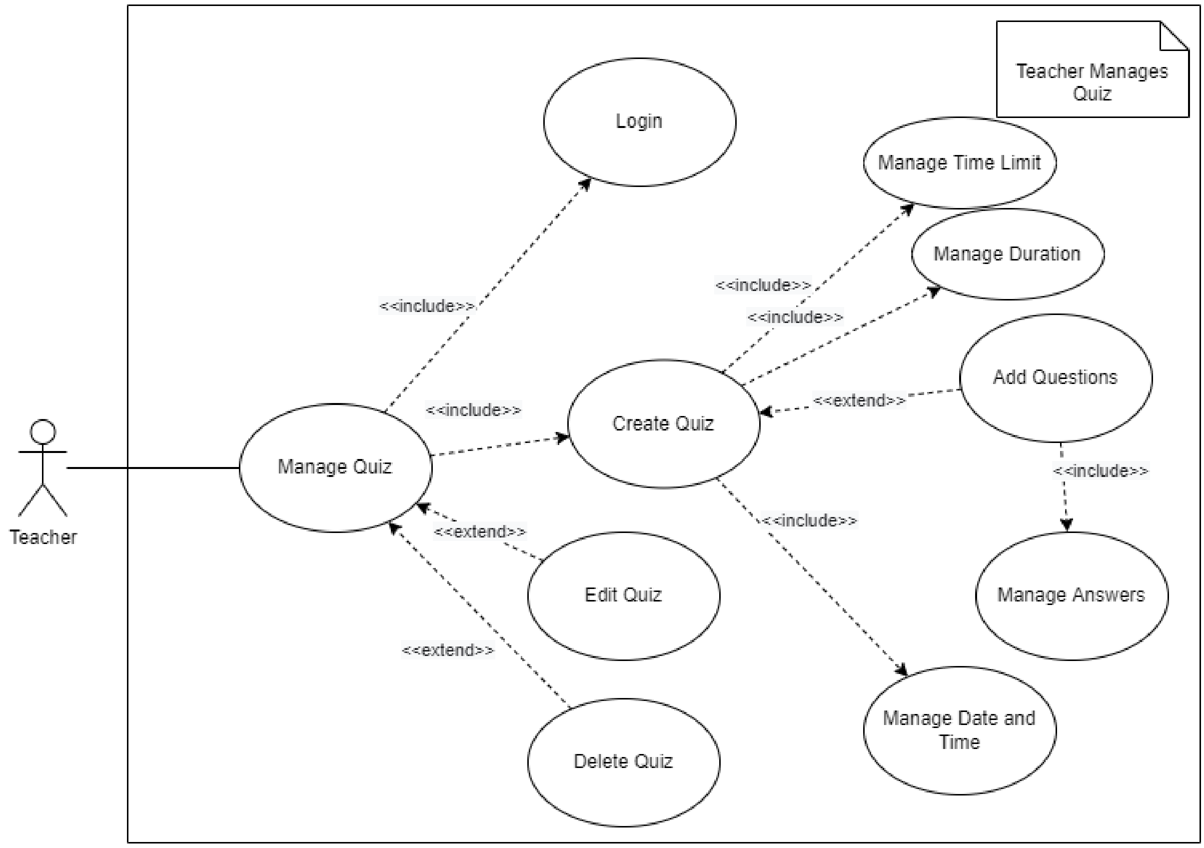


Figure III.8: Use Case: Teacher manages Quiz

Use Case	Teacher Manages Quiz
Description	Teacher manages a quiz
Actors	1. Teacher
Preconditions	1. Teacher must be logged in 2. Teacher must have the permission to manage quizzes 3. In case of modifying questions, answers and KCs, the quiz must not be active yet.
Main Scenario	1. Teacher Accesses Dashboard 2. Teacher Clicks Create New Quiz in Course Navigation Bar 3. Teacher fills necessary information to create a new quiz 4. Teacher is redirected to Manage Quiz Questions Page 5. Teacher Adds a new Question 6. Teacher Adds a few Answers to the Question 7. Teacher marks answers as correct or incorrect 8. Teacher Adds Knowledge Components to the Question 9. Teacher Saves the changes 10. Teacher notifies students of the newly created quiz.
Alternative Scenario	8. Teacher doesn't find the KC he wants to add to the question 6. Teacher is prompted if he wants to create the new KC 7. Teacher wants to create the KC, he is redirected to the "manage Course KC Page" 8. Teacher creates new KC. 9. Teacher is redirected back to Manage Quiz Questions Page. 10. Teacher adds the newly created KC to the question. 11. Teacher saves the changes
Post-conditions	1. Quiz must have at least one question. 2. Every Question must have at least 2 answers. 3. Every Question must have at least one correct answer. 4. Every Question must have at least one wrong answer.

Table III.5: Use Case: Teacher Manages Quiz

Teacher Monitors Quiz

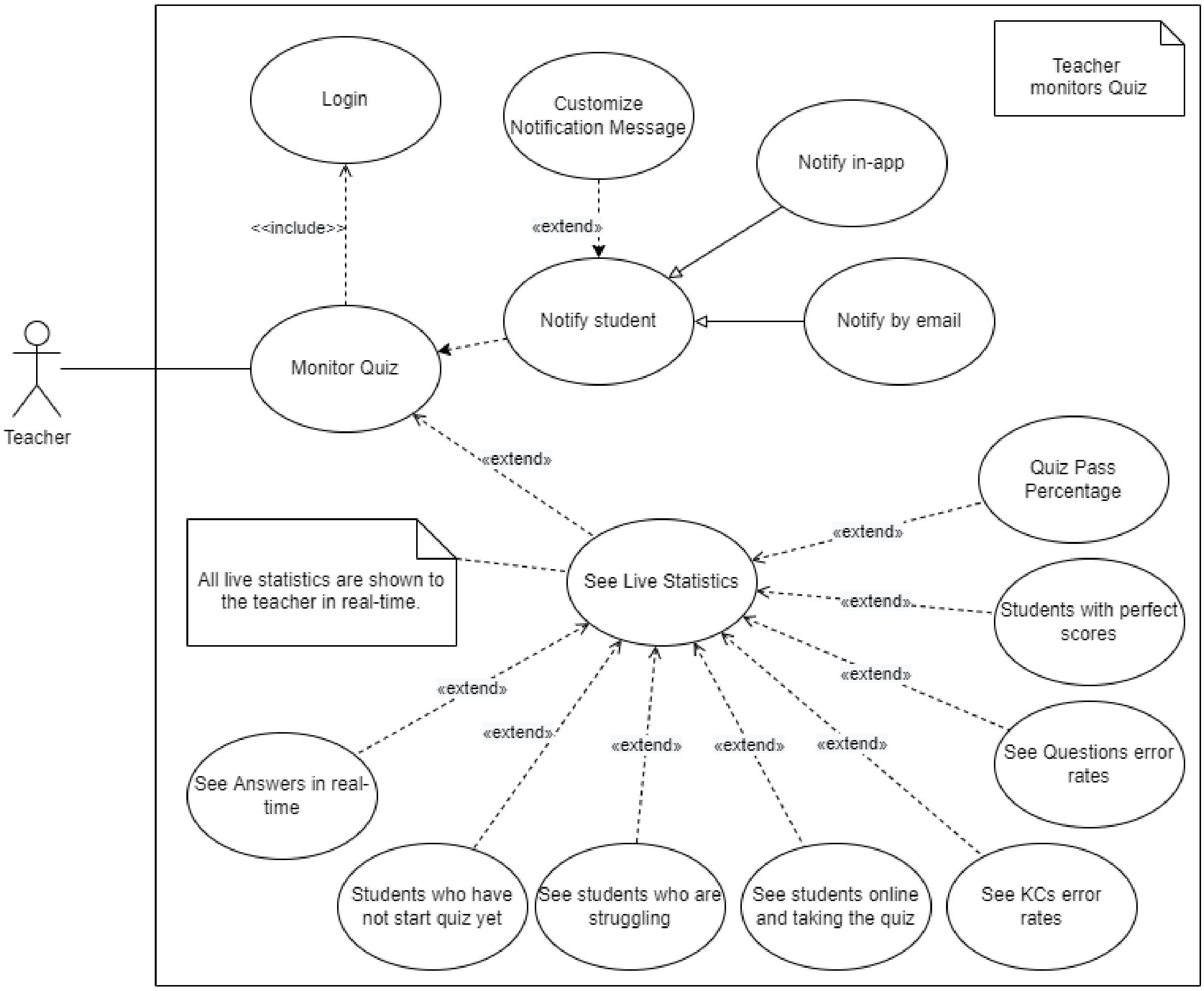


Figure III.9: Use Case: Teacher monitors Quiz

Use Case	**Quiz System**
Description	Teacher monitors a quiz
Actors	1. Teacher 2. Student
Preconditions	1. Teacher and Students must be logged in 2. Each Questions must have its KCs 3. Each KC must have a relationship with at least one Course Lesson.

Main Scenario	1. Quiz Starts 2. Students are notified 3. Students start Quiz 4. Teacher sees a list of students who are attempting the quiz 5. Students responds to question incorrectly 6. Teacher sees that student answered a question with a wrong answer 7. Teacher sees which KC his student is struggling in
Alternative Scenario	5. Students Answer two questions that share the same KC, one correctly and the other not. 6. Teacher is notified that these question probably do not belong to the same KC 7. Teacher goes to "Manage KCs" page to separate KC.
Post-conditions	1. All question responses are recorded 2. Every student who submits his quiz is recorded 3. Students who didn't submit are recorded separately

Table III.6: Use Case: Teacher Monitors Quiz

Monitor Quiz Interaction in Sequence Diagram

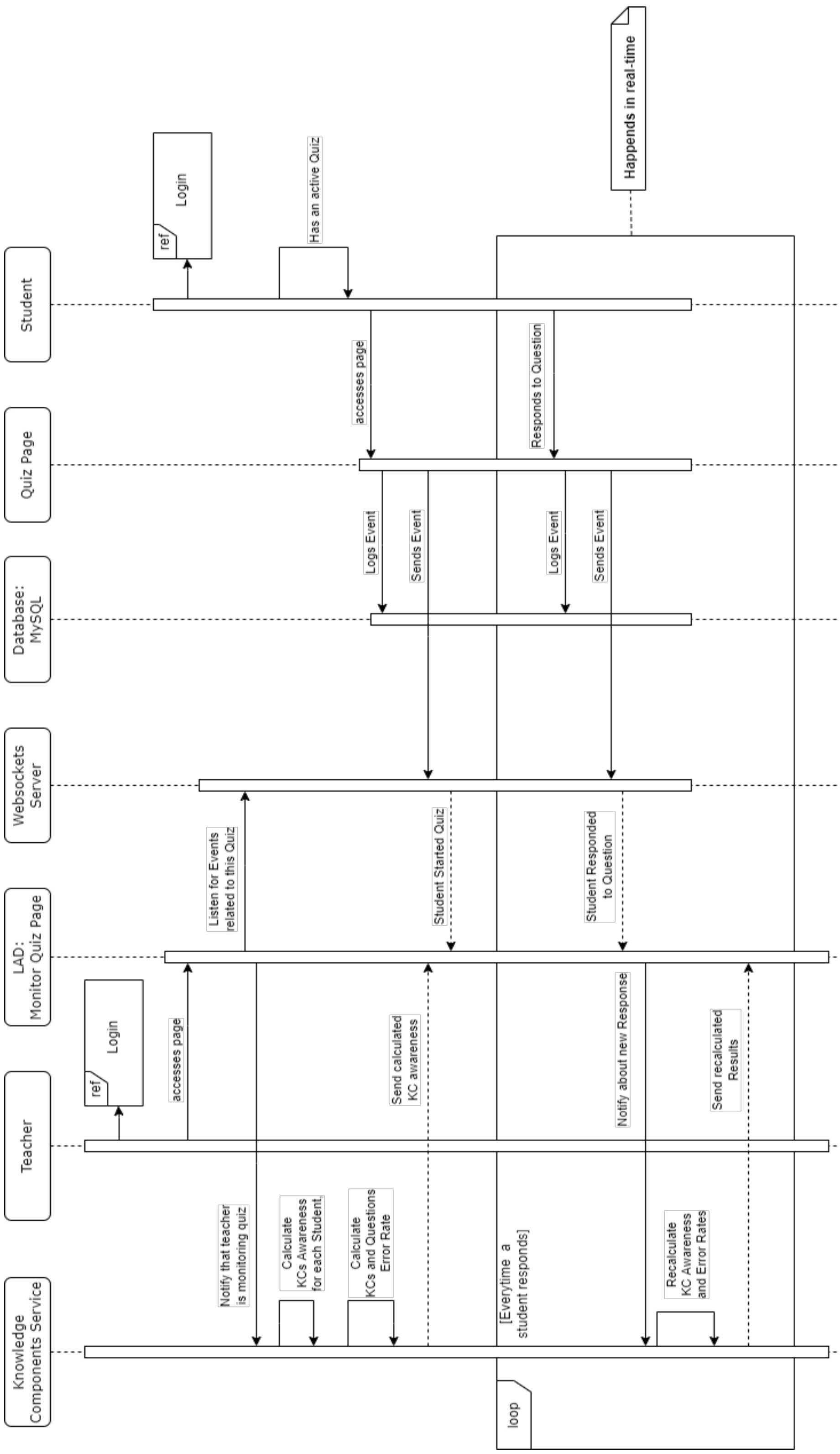


Figure III.10: Sequence: Teacher Monitors Quiz

Student Takes Quiz

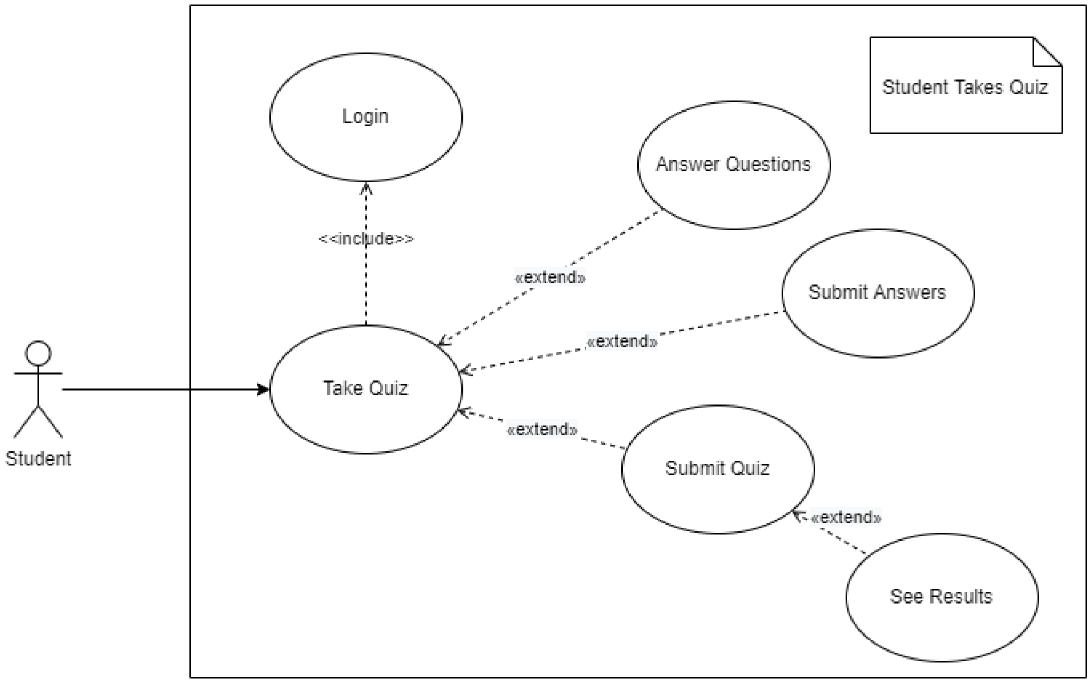


Figure III.11: Use Case: Student takes Quiz

Use Case	**Student Takes Quiz**
Description	When a student's course has a quiz, a student is notified of the start date, when the date arrives, a student can access the quiz page and respond to the questions
Actors	1. Student
Preconditions	1. Student must be logged in 2. Student must have permission to access Quiz 3. Quiz must be in active state
Main Scenario	1. Student accesses his courses page 2. Student accesses course "Software Engineering" 3. Student visits quizzes page of this course 4. Student sees that he has an active quiz 5. Student starts the quiz 6. Student answers all multi-choice questions. 7. Student submits his quiz. 8. Student is redirected to the results page. 9. Student can see all the questions with his responses and the correct answers

Alternative Scenario	6. Student answers some questions 7. Quiz duration is over 8. Student is kicked off from the quiz page into the results page. 9. Student can see all the questions with his responses and the correct answers
Post-conditions	1. Teacher is notified when student submits Quiz.

Table III.7: Use Case: Teacher Manages Knowledge Components

Knowledge Components System

Teacher Manages Knowledge Components

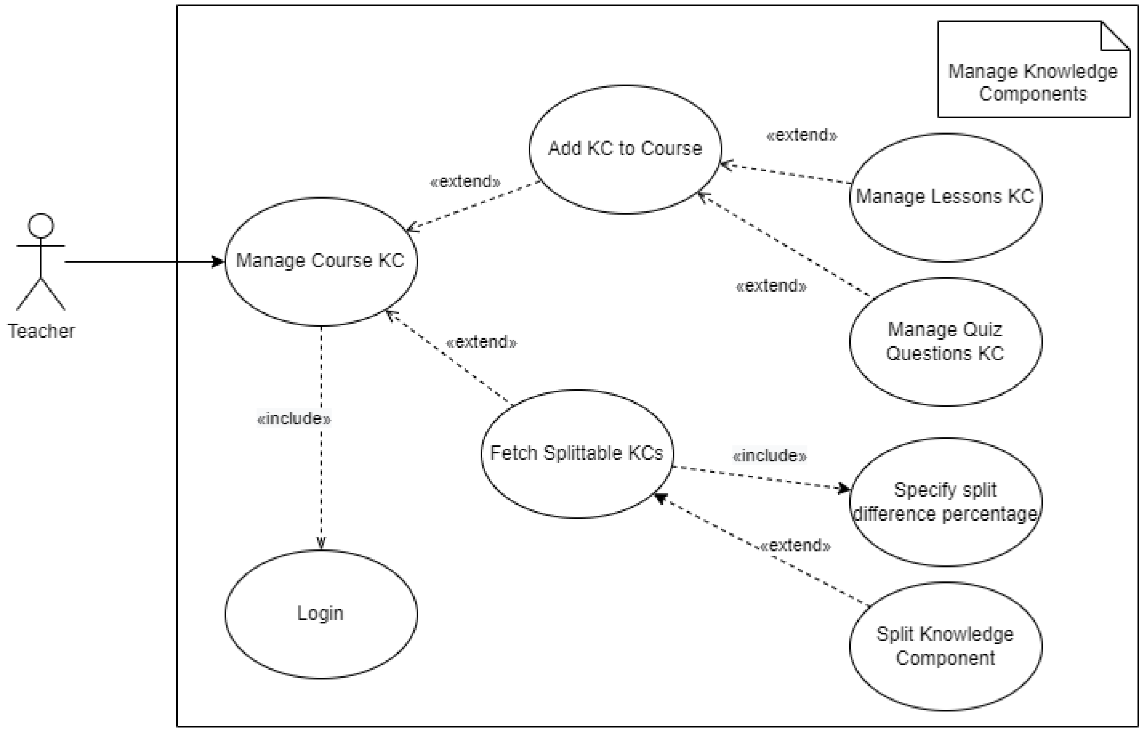


Figure III.12: Use Case: Teacher Manages Knowledge Components

Use Case	**Teacher Manages Knowledge Components**
Description	Teacher can manage his course and course materials' KCs.
Actors	1. Teacher
Preconditions	1. Teacher must be logged in 2. Teacher must have the permission to manage KCs 3. KCs cannot be edited if they belong to certain quiz questions and said quiz is active

Main Scenario	1. Teacher Accesses Dashboard 2. Teacher accesses a finished Quiz's monitor page3. Teacher goes to the "Questions Error Rate Bar Chart" 4. Teacher notices that two questions that share the same KC, have a considerable difference in Error Rate5. Teacher is prompted to split this KC. 6. Teacher clicks on Split KC button. 7. Teacher is redirected to "Split KC page" 8. Teacher fills the necessary information for splitting KC. 9. Teacher Saves the changes
Alternative Scenario	3. Teacher goes to the "KCs Error Rate Bar Chart" 4. Teacher notes all KCs that have a high Error Rate 5. Teacher selects those KCs 6. Teacher notifies all students that need to revise the lessons related to these KCs
Post-conditions	1. Created KC's name must be unique.

Table III.8: Use Case: Teacher Manages Knowledge Components

Teacher Monitors Knowledge Components

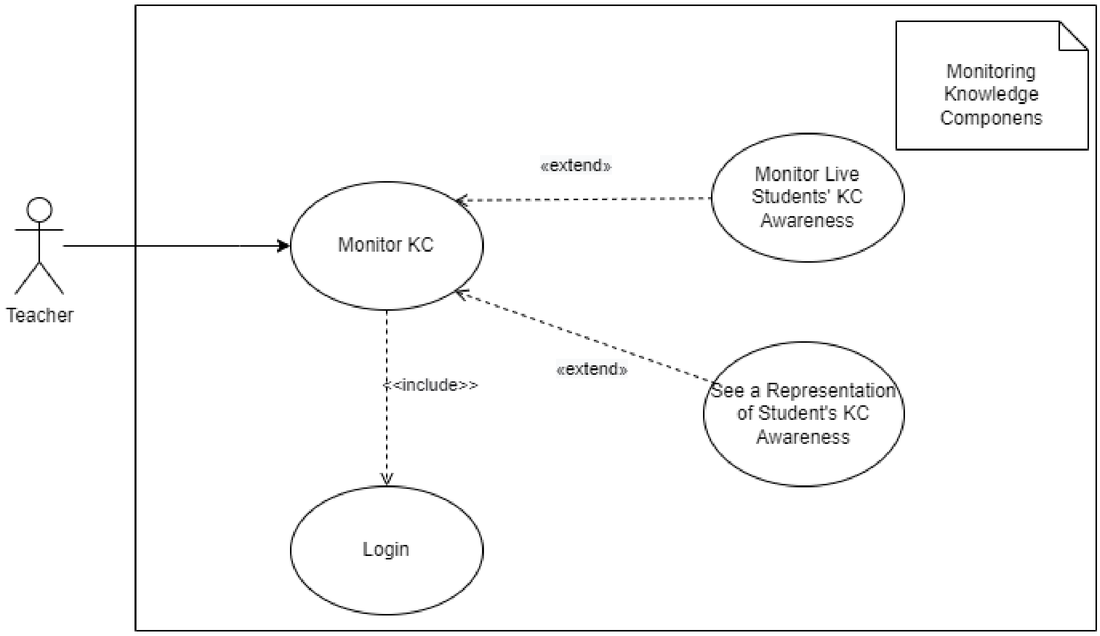


Figure III.13: Use Case: Teacher Monitors Knowledge Components

Use Case	**Teacher Monitors Knowledge Components**
Description	When (or after) students are taking a quiz, a teacher can monitor their awareness of KCs.

Actors	1. Teacher 2. Student
Preconditions	1. Teacher and Student must be logged in 2. Teacher must have the permission to monitor Quiz 3. Student must have the permission to take the Quiz
Main Scenario	1. Student Starts a Quiz 2. Student answers some question 3. Teacher Accesses "Quiz Monitor" page 3. Teacher is notified every time a student takes action in the quiz page. 4. Teacher notices that this student is barely answering the questions correctly. 5. Teacher sees which KCs his student is struggling in. 6. Teacher accesses "Monitor Student in Quiz" page. 7. Teacher sees a Radar Chart representation of the quiz's KCs, and the student's awareness of these KCs. 8. Quiz time is over 9. Teacher notifies student of the lessons that are related to the KCs he struggled in.
Alternative Scenario	4. Teacher notices that all students are struggling in the same KC 5. Teacher can see all the questions and lessons related to this KC. 6. Teacher notifies students that these Lessons must be revised.
Post-conditions	1. Every student must have their KCs score and awareness for this quiz calculated.

Table III.9: Use Case: Teacher Manages Knowledge Components

III.4 Application Architecture

III.4.1 Class Diagram

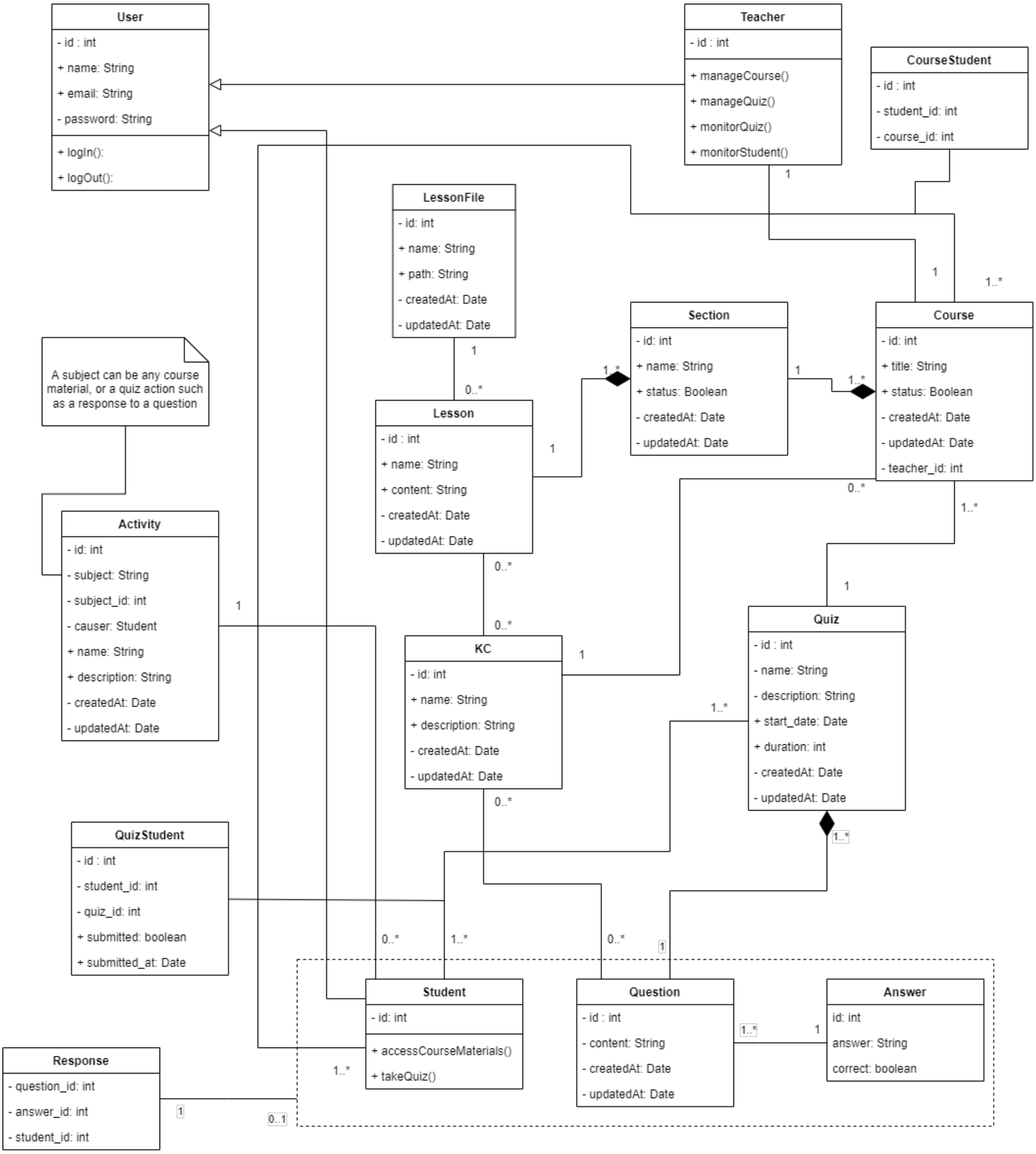


Figure III.14: Class Diagram

III.5 Conclusion

In this chapter, we have proceeded to a detailed conceptual study based on a variety of UML diagrams. This allowed us to have both a broad and detailed vision of the architecture of our application and helped us conceptualize it in order to make it a reality. Such a milestone will be primordial to the next step of our document, since the implementation phase will have a clear pathway to follow which will be the topic of our upcoming chapter.

Chapter IV

Implementation

IV.1 Introduction

In this chapter, we will go into details about the steps it took to build our project. We will first discuss our Event-Logging solution, then go over the Knowledge Components service, the cite the several tools we used. Next, we will proceed to the description of the major interfaces illustrated by screenshots, and finally, we will talk about the deployment process of our Dashboard and Websockets Server.

IV.2 Logging Students Activities

IV.2.1 Setup

To successfully log a student's activities, we used middlewares, which are convenient mechanism for inspecting and filtering HTTP requests entering a web application. In most websites, they are used for Authorization, the Encryption or Decryption of cookies, and a great way to limit the number of requests a user can make in a certain period.

In our case, it is a convenient method to log students activities, as shown in the figure below. All the routes concerning a student are grouped together, this way, every time a student makes a request in his dashboard, it passes by this Route group that is prefixed with the word "student".

lad - web.php

```php
Route::prefix('student')
    ->name('student.')
    ->middleware([ 'auth', 'role:student', 'student_last_activity', 'log' ])
    ->scopeBindings()
    ->group(function () {

        // All students routes grouped ..
```

Figure IV.1: Code: Student Routes Middleware

In Laravel web routes file, calling the method "middleware" on a route group tells PHP to assign every middleware to every route in that group, and to fire their handle method before every request. The figure above shows four easily identified middlwares:

1. The *'auth'* middleware: checks if the user is authenticated.
2. The *'role:student'* middleware: as the name indicates, check if the authenticated user has a role of "student".
3. The *'student last activity'* middleware: records the student's last activity and caches his online status for 30 seconds.
4. The *'log'* middleware: logs the student's actions in the dashboard.

Since the release of MySQL v5, the open-source relational database management system (DBMS) has shown a good performance increase that is close to 3000 transactions per second for multi-threading and more than 150 transactions for applications with a single-thread [1]. This improvement allows a web application dedicated for Learning Analytics and course management to successfuly use it for logging student's activity without any hindrance to the performance of the dashboard and the user experience (UX).

When tested in our application, every web page that heavily interacts with the database (Mainly the Quiz Monitor Page), shows that all the database queries only take about a 100ms in a local environment, and less than a second in a free shared-hosting environment.

IV.2.2 Sending Logs to the LAD in Real-Time

There are several HTTP/1.1 and HTTP/2 solutions that allow web applications to communicate with the server bidirectionally including server-pushed messages, streams and multiplexing and all types of HTTP Polling [2]. For our case, the 'server-pushed messages' [3] solution is unnecessary (even overkill) as it is more suited for allowing Chat applications to broadcast new chat messages to all the clients in a chat room that have multiple open Transmission Control Protocol (TCP) connections. And after testing Long Polling, it became obvious that a component based dashboard web page that heavily relies on Real-Time interactions with the server becomes really slow and breaks all user experience.

That is why, because of its perfect coupling with the Event-Based System (The Observer Pattern) that is our Real-Time Dashboard, we chose the Websockets solution over the others. A great way to illustrate how every request made by the student is logged, transferred into an event and sent into the LAD using websockets can found in the "Monitor Quiz" Sequence Diagram III.10.

Events serve as a great way to decouple various aspects of a web application, since a single event can have multiple listeners that do not depend on each other `https://laravel.com/docs/9.x/events#introduction`. In our case, listening for them in the dashboard is an easy and a well organized task when working on a component based front-end, the figure below shows how the browser receives the events that are triggered whenever a student answers a quiz question, the event carries a small payload that doesn't exceed a few kilobytes that makes listening for events a light task in contrast to the Long Polling solution, the payload in the case of the event mentioned above, only consists of the identifiers of the student, plus the question, answer, response, and whether the response is correct or not. [4]

[1] `https://blog.dbi-services.com/mysql-versions-performance-comparison/`
[2] `https://www.educative.io/edpresso/what-is-http-long-polling`
[3] `https://building.lang.ai/our-journey-from-websockets-to-http-2-4d069c54effd`

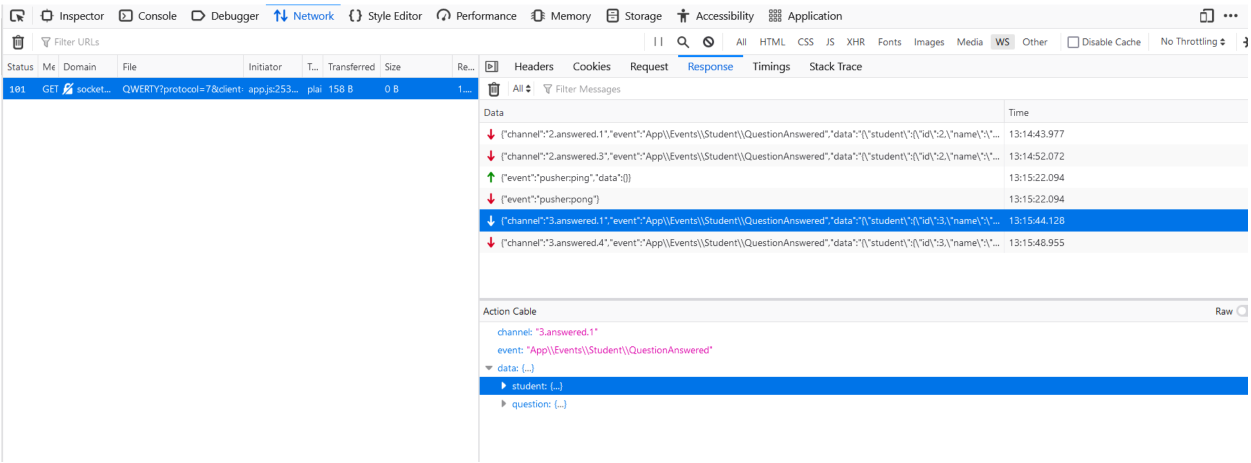


Figure IV.2: Quiz Response Events coming to Quiz Monitor Page via a Websockets Server

That is the case for an event that travels directly to the front-end and doesn't need to run by any logic in between, as for the case when logic needs to be executed, listening for Events happens on the server, so that we can call the *KC Service* that is decoupled from the Model-View-Controller (MVC) architecture that governs our dashboard, and the Observer Pattern that takes care of all events.

IV.3 Tools Used

IV.3.1 PHP

PHP is an open-source server-side scripting language, and a powerful tool that is especially suited to web development [5]. It is well documented and has a good following and community. With the release of PHP 8, the addition of JIT compilation has become a big plus in using PHP.

IV.3.2 Laravel

Laravel is an open-source PHP framework, It follows a model-view-controller design pattern. Laravel reuses the existing components of different frameworks (Symfony and Doctrine, GuzzleHTTP, etc.) which help in creating a web application. The web application thus designed is more structured and pragmatic. [6]

IV.3.3 Apache

Apache is a free and open-source software that allows users to deploy their websites on the internet. It is one of the oldest and most reliable web server software maintained by the Apache Software Foundation.[7]

[5] https://www.w3schools.com/php/
[6] https://www.tutorialspoint.com/laravel/laravel_overview.htm
[7] https://www.hostinger.com/tutorials/what-is-apache

It's fast, reliable, and secure. And Apache can be highly customized to meet the needs of many different environments by using extensions and modules.[8]

IV.3.4 MySQL

MySQL, the most popular Open Source SQL database management system, is developed, distributed, and supported by Oracle Corporation [9]. It provides great performance even for single-threaded web applications that need to log user's interaction and his activities.

w

IV.3.5 Laragon

Laragon is a portable, isolated, fast and powerful universal development environment for PHP, Node.js, Python, Java, Go, Ruby. It is fast, lightweight, easy-to-use and easy-to-extend.

Laragon is great for building and managing modern web applications and the versions of dependencies (PHP, MySQL, etc.). It mainly serves as a replacement to tools like Docker and Kubernetes for small teams and small applications.

IV.3.6 WebSockets

The WebSockets API is an advanced technology that makes it possible to open a bi-directional communication session between the user's browser and a server.[10] This API will allow us to can send messages to a server and receive event-driven responses without having to poll the server for a reply.

[8] https://www.wpbeginner.com/glossary/apache/
[9] https://dev.mysql.com/doc/refman/8.0/en/what-is-mysql.html
[10] https://developer.mozilla.org/en-US/docs/Web/API/WebSockets_API

IV.4 Application User Interfaces

IV.4.1 Login Page

This page is the first page the user sees, the user is asked to enter his address and password, then he is redirected to his home page, depending on his role (student / teacher).

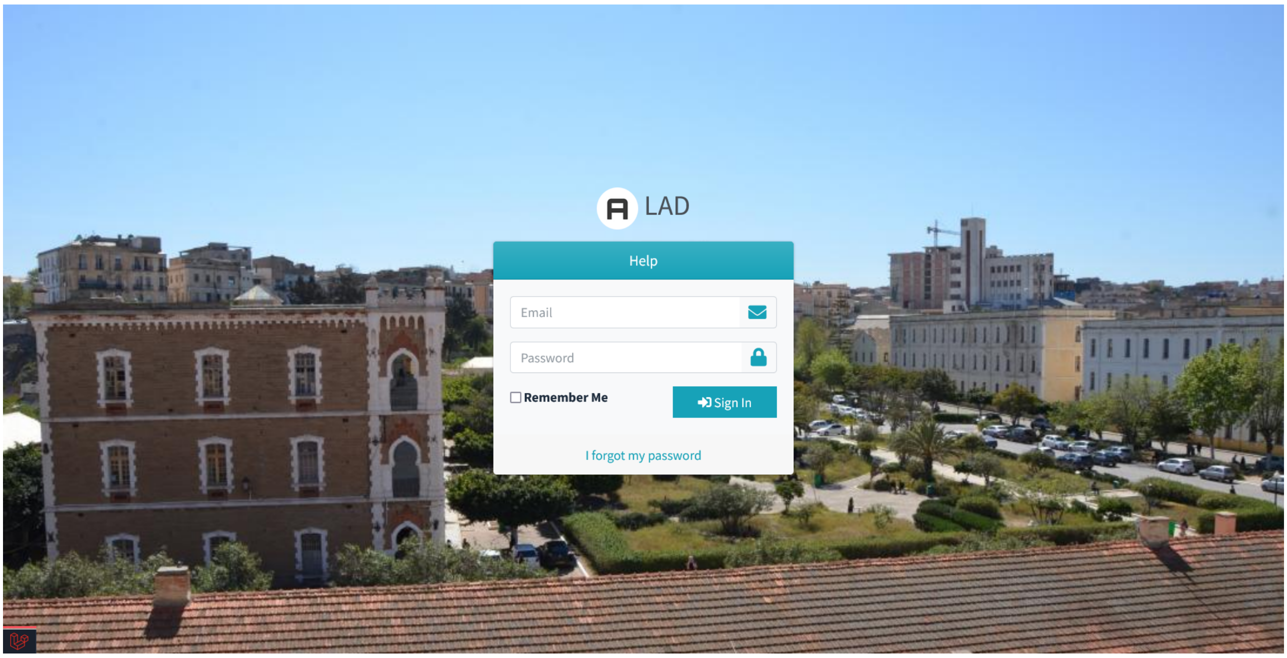


Figure IV.3: Authentication interface

.

From now on, we will divide this presentation according to the two major actors of the system, namely: Student and teacher.

IV.4.2 Student User Interfaces

Once the student is logged in, he's able to view his courses, sections and lessons and is notified by the newly added ones as shown in the figures below.

Course Page

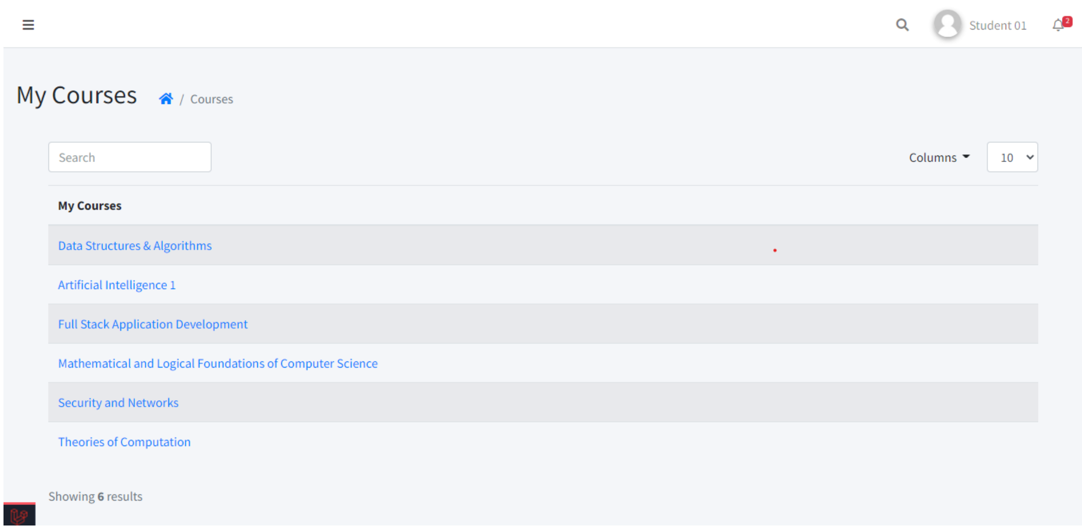


Figure IV.4: Course page interface

.

Section Page

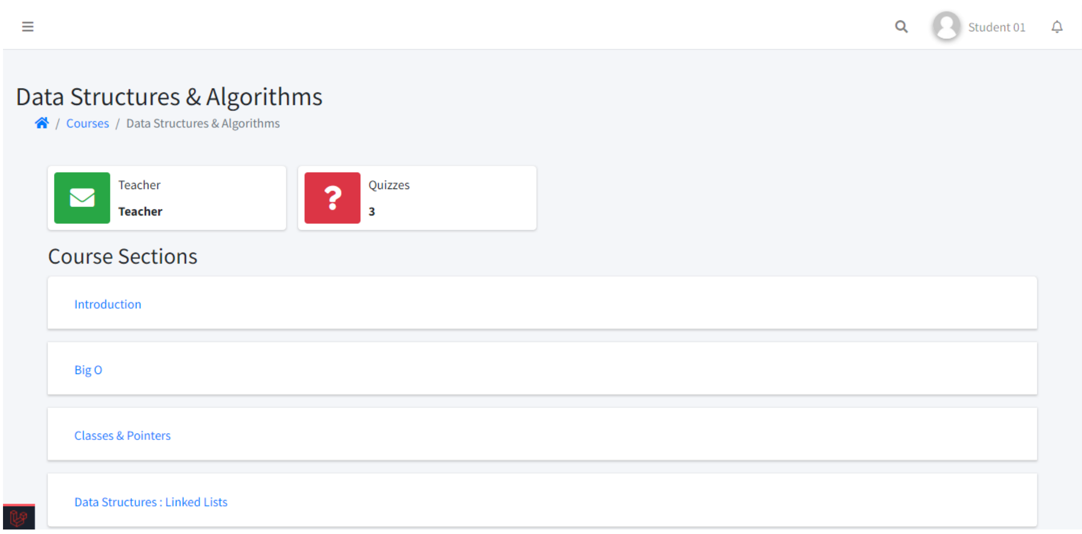


Figure IV.5: Section page interface

.

Lesson Page

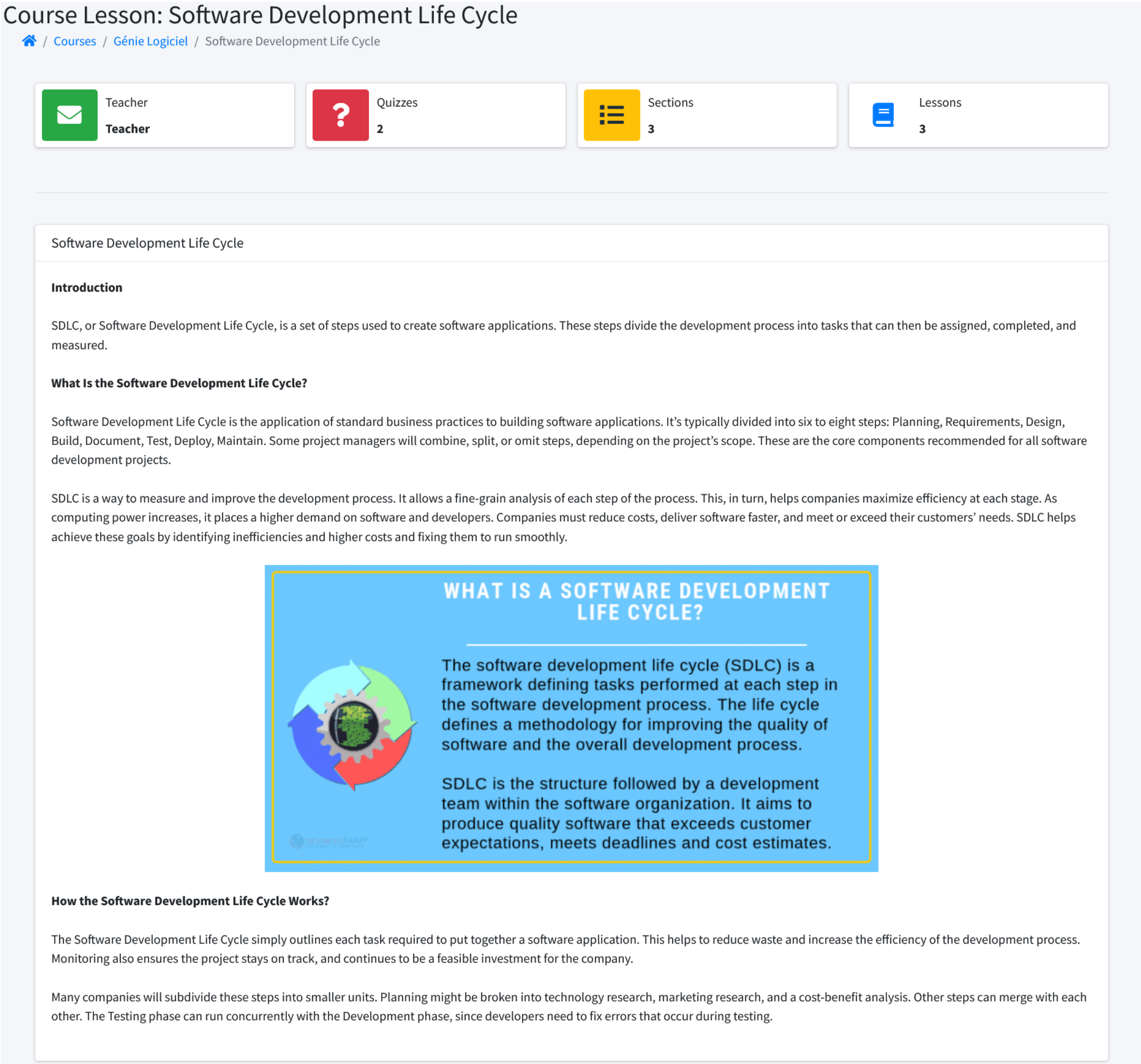


Figure IV.6: Lesson page interface

.

List of quizzes Page

In addition, he can see the list of all quizzes and their status (active, upcoming), take one and see his results.

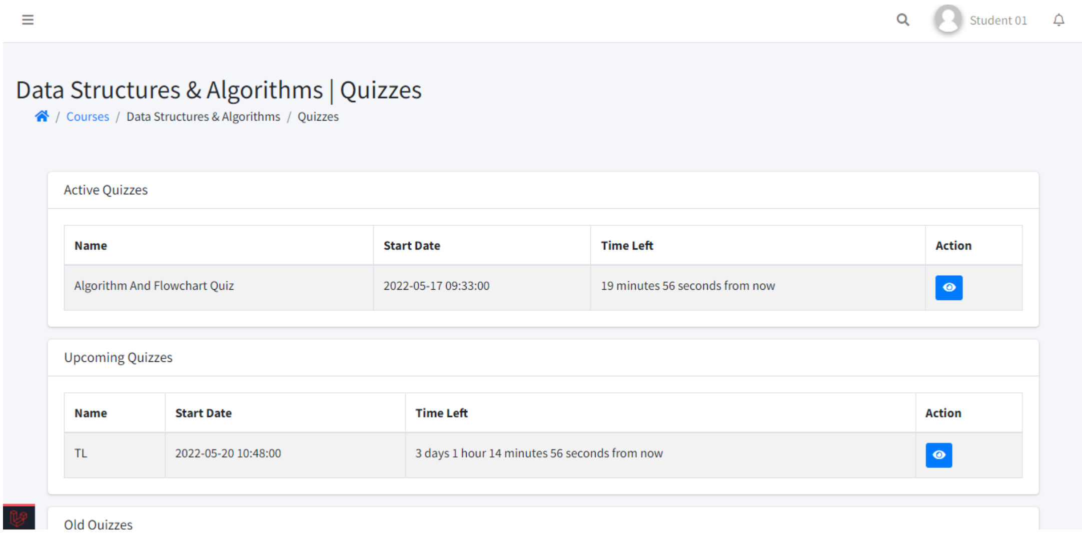


Figure IV.7: List of quizzes page interface

.

Quiz Page

When the student starts the quiz a countdown timer is shown, so he can manage his time properly.

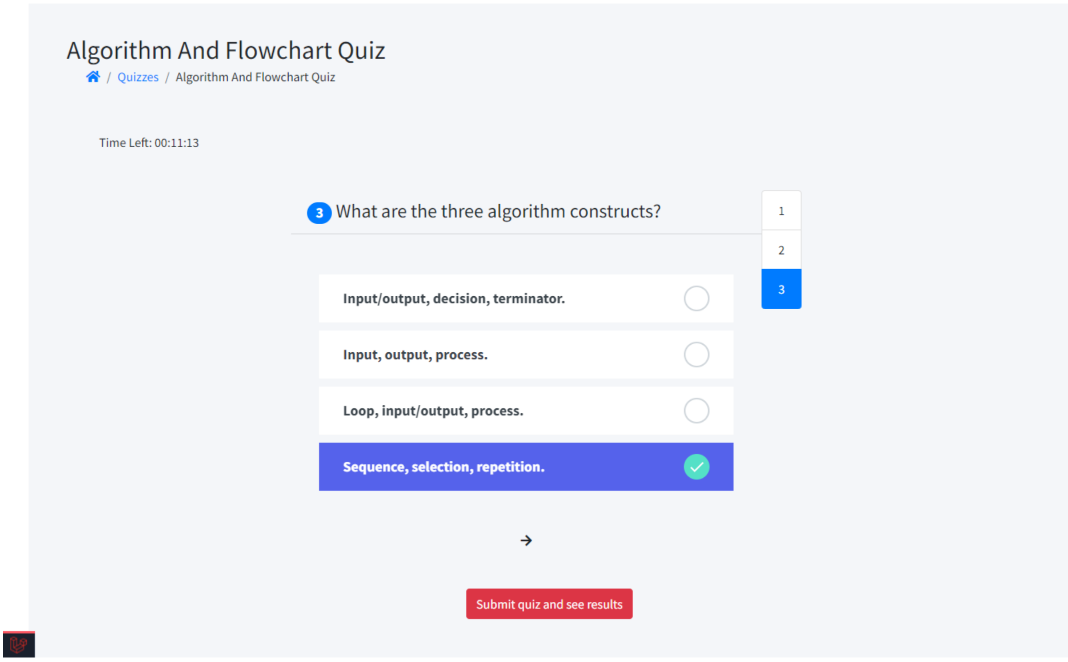


Figure IV.8: Quiz page interface

.

Quiz Results Page

Once he finishes he can check his results and see where he is correct or not.

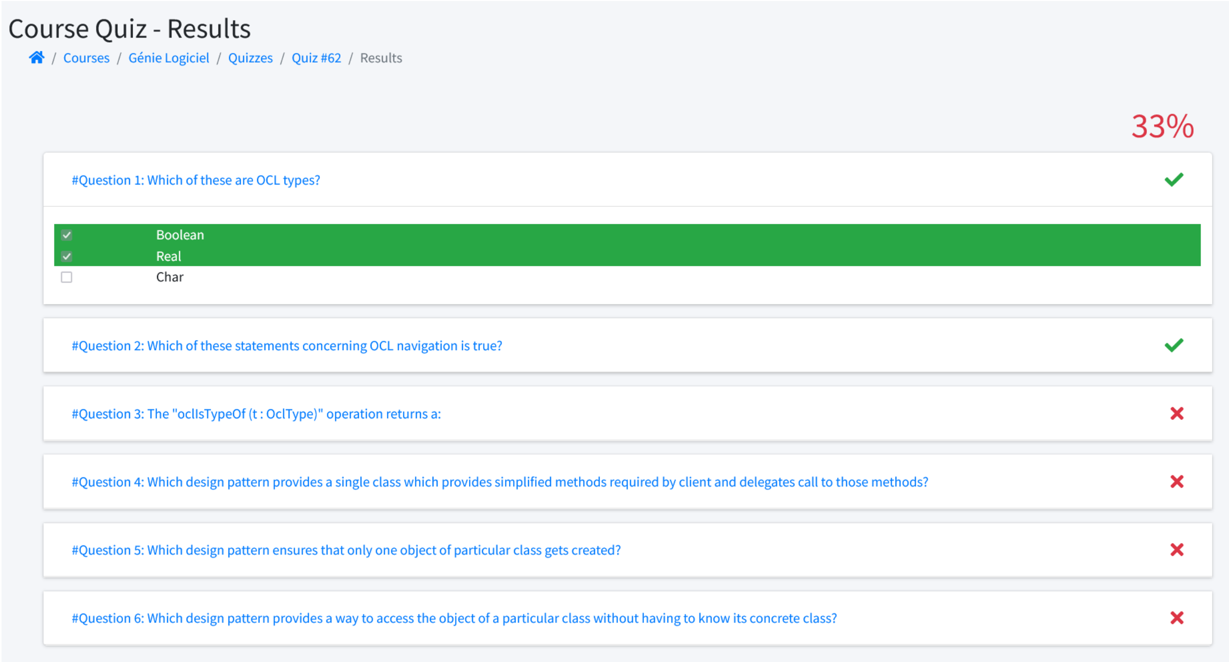


Figure IV.9: Quiz Results interface

.

IV.4.3 Teacher User Interfaces

Course Monitor Page

Once the teacher is authenticated, he has access to a student statistics overview page in which he sees the online students and a bar chart showing their visits to the course, combined with a line chart with their total visits to the course lessons.

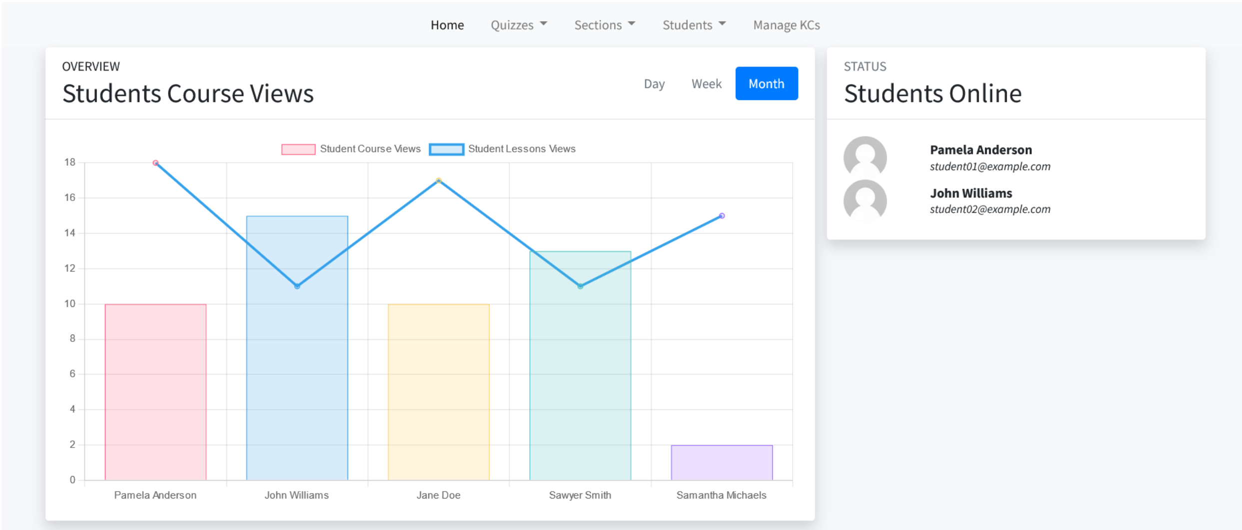
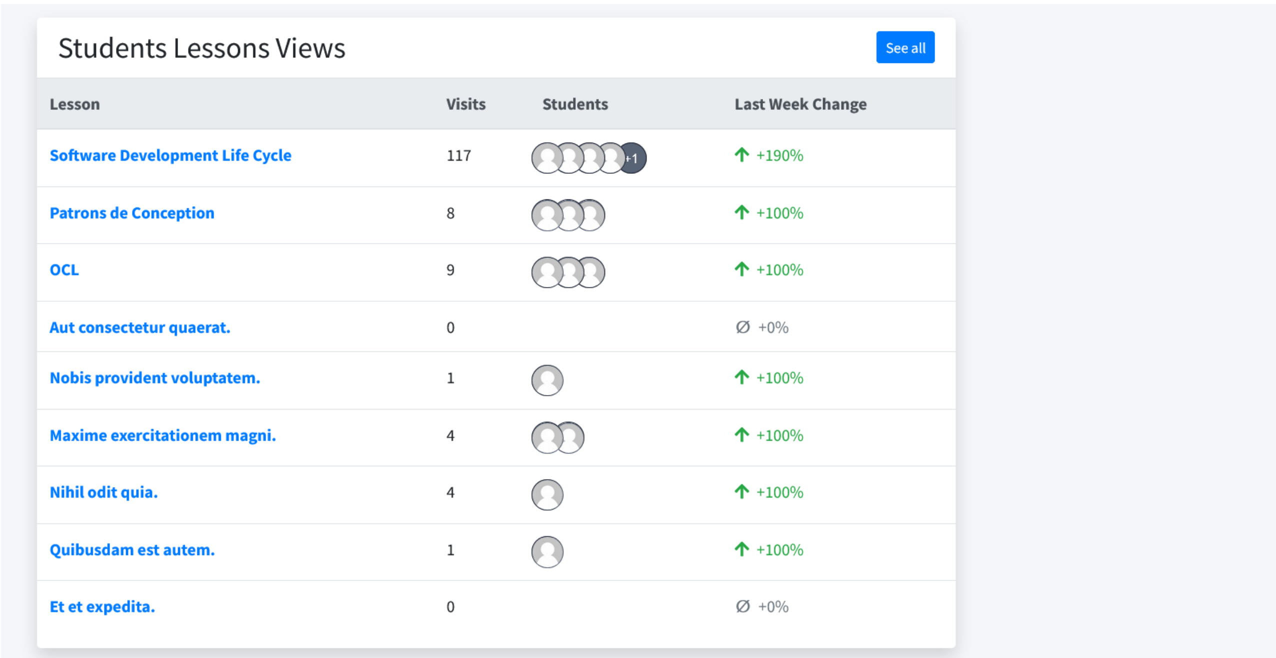


Figure IV.10: Overview: Course Monitor Page

After scrolling down the same page, the teacher will see a list of lessons and the students who visited these lessons, with a percentage change from last week to compare visits.



Lesson	Visits	Students	Last Week Change
Software Development Life Cycle	117	+1	↑ +190%
Patrons de Conception	8		↑ +100%
OCL	9		↑ +100%
Aut consectetur quaerat.	0		Ø +0%
Nobis provident voluptatem.	1		↑ +100%
Maxime exercitationem magni.	4		↑ +100%
Nihil odit quia.	4		↑ +100%
Quibusdam est autem.	1		↑ +100%
Et et expedita.	0		Ø +0%

Figure IV.11: Course Lessons Views

.

Section Page

A teacher manages his course's sections and lessons making sure to assign the relevant Knowledge Components to each lesson.

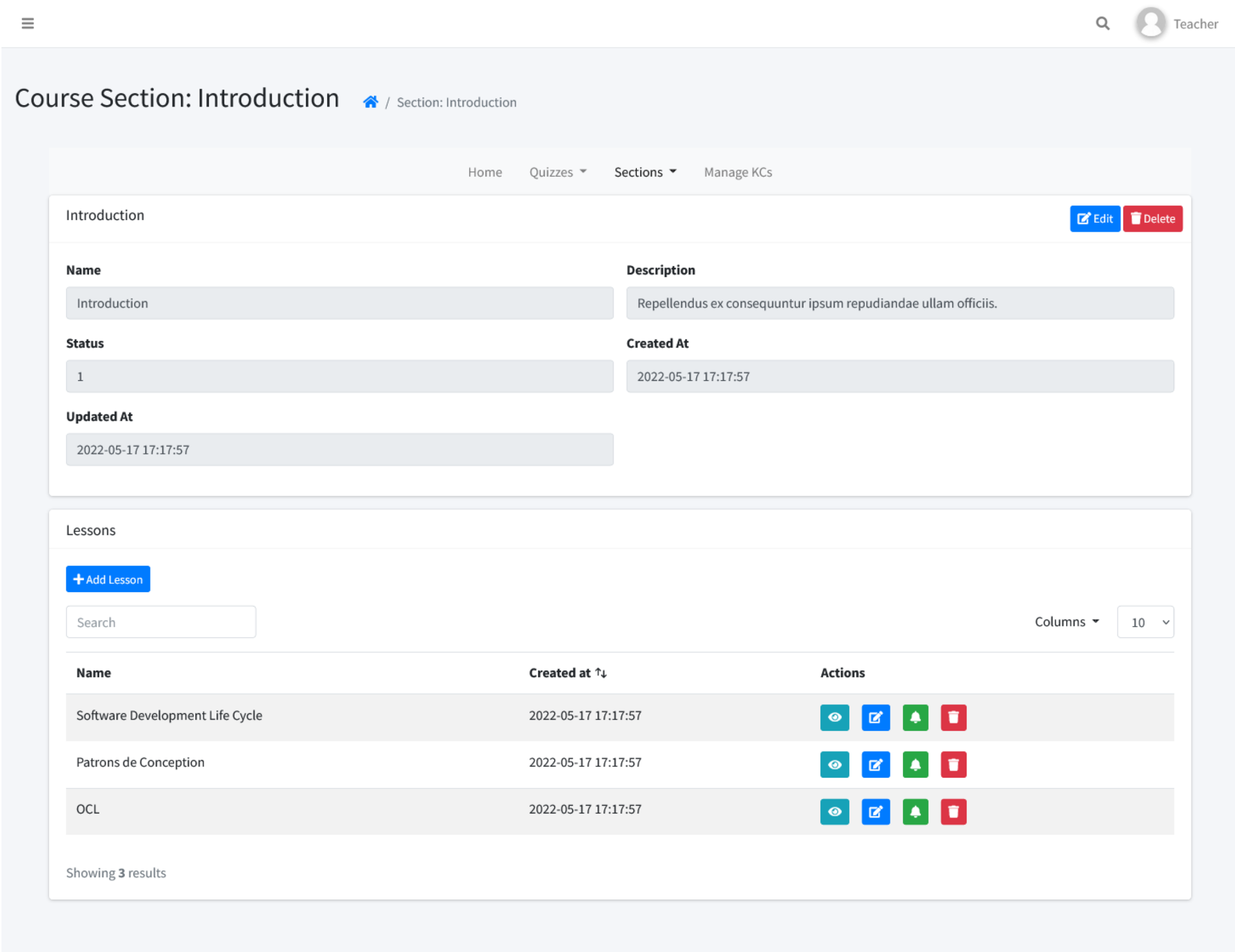


Figure IV.12: Section Page

.

Lesson Edit Page

The teacher can modify his lessons according to his preferences.

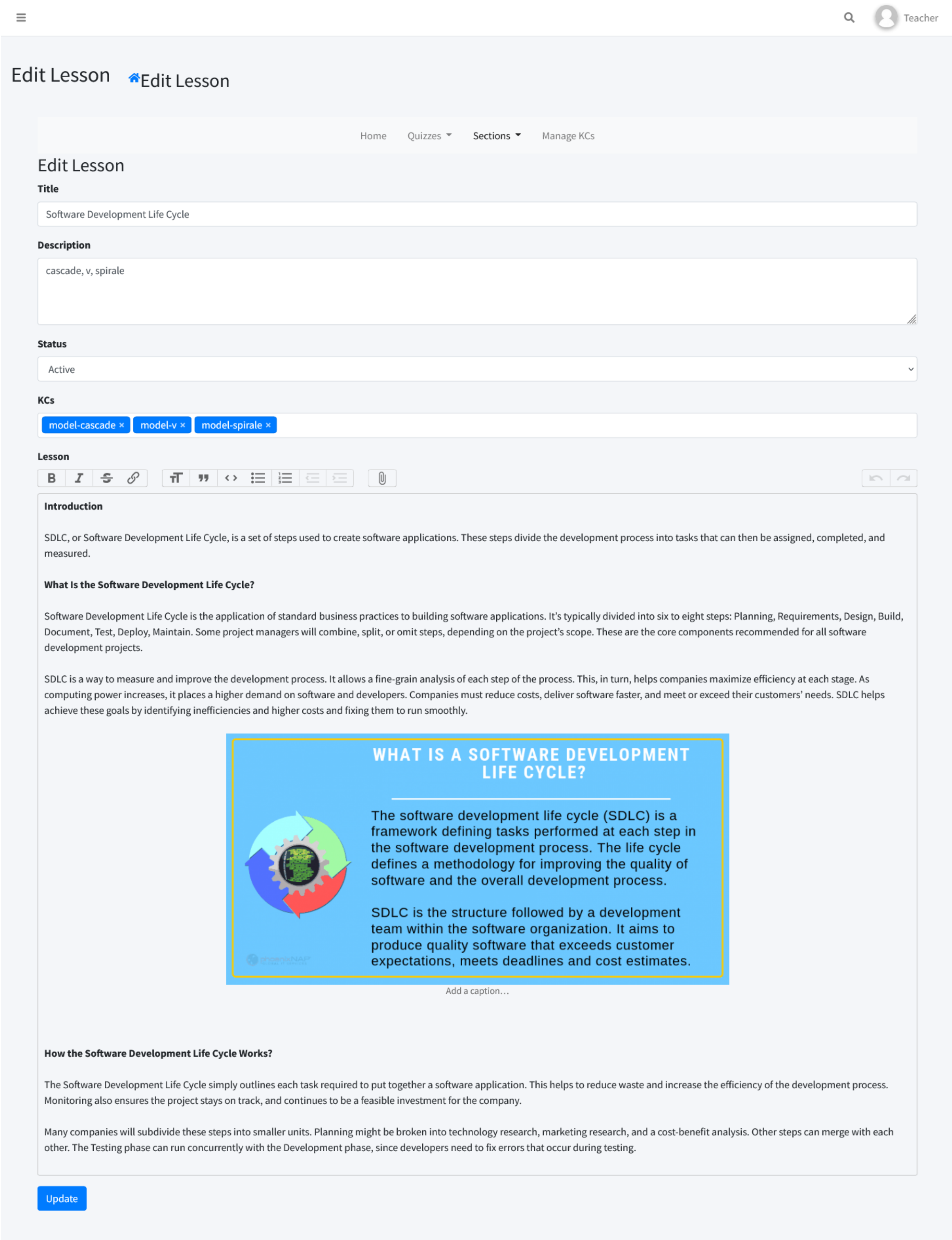


Figure IV.13: Edit Lesson Page

Quiz Edit Page

While the teacher is managing a quiz, he ensures a KC for each quiz's question, he can also reorder quiz questions, and reschedule the quiz start date, and notify students either via email or in-app notifications.

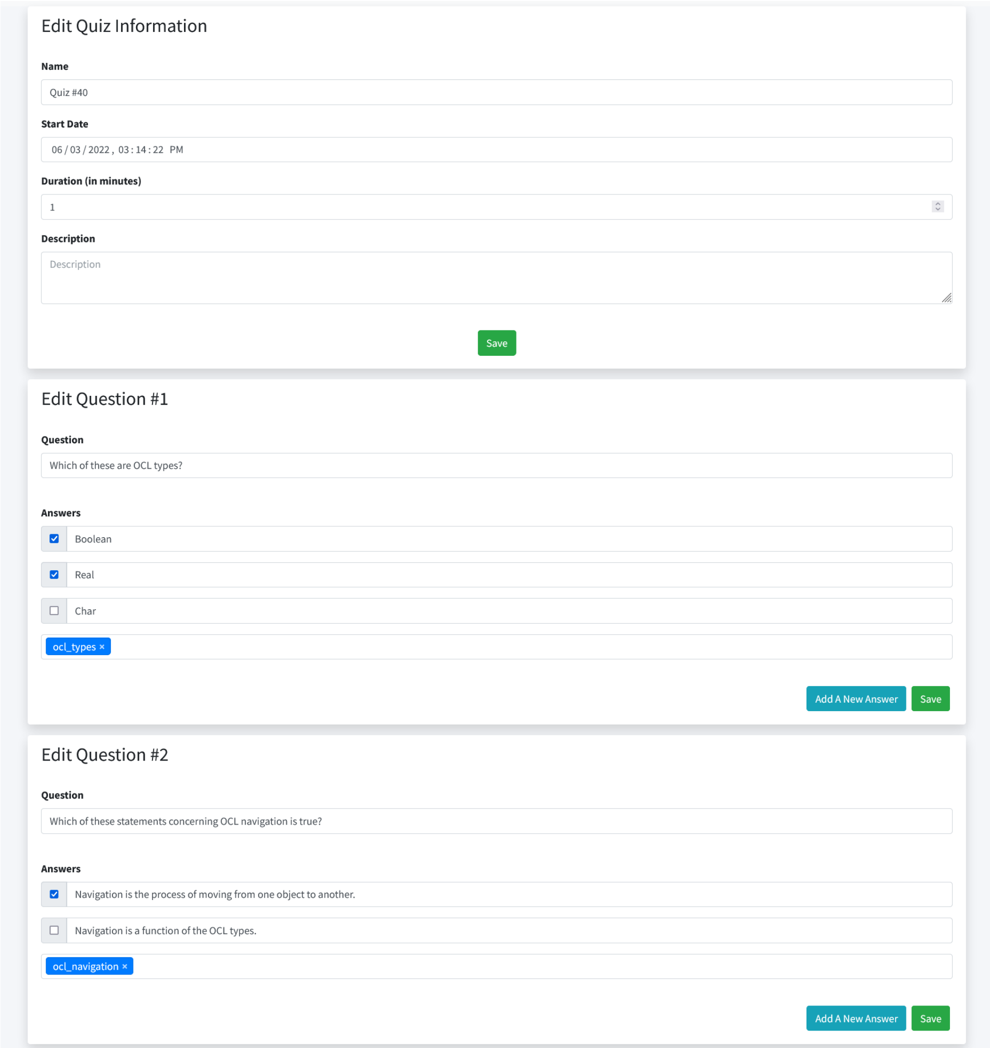


Figure IV.14: Edit Quiz Page

.

Course Knowledge Components Manage Page

The teacher has a page where he can manage his course Knowledge Components, it is divided into two sections: Main Section

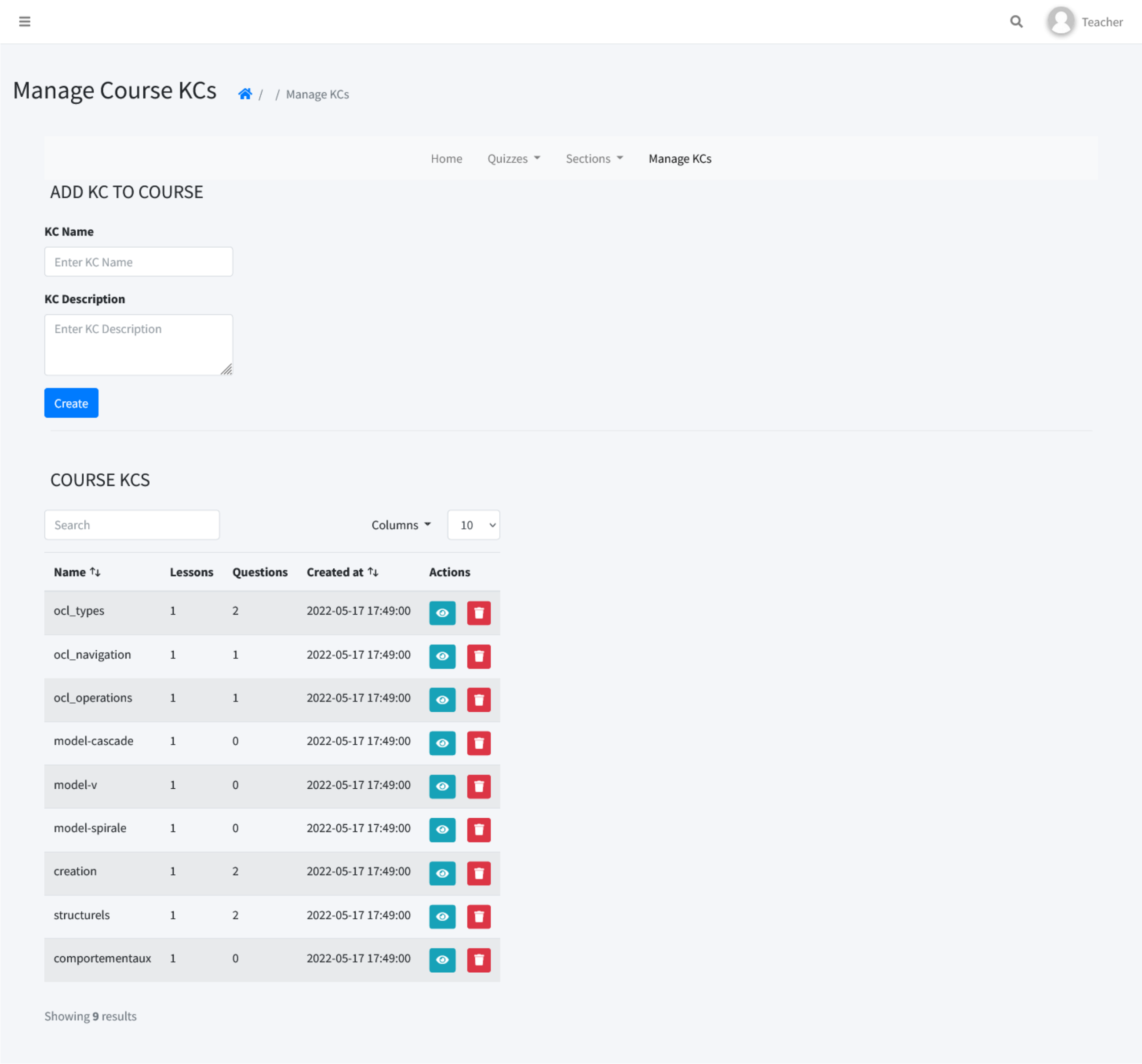


Figure IV.15: Manage All Course KCs Page Section

.

After a course has collected several quiz results throughout a semester, a teacher is able to take advantage of the collected data to improve his course, by fetching possible splittable KCs according to quiz questions grouped by similarity in KCs, and the difference of error rates, specified by the teacher.

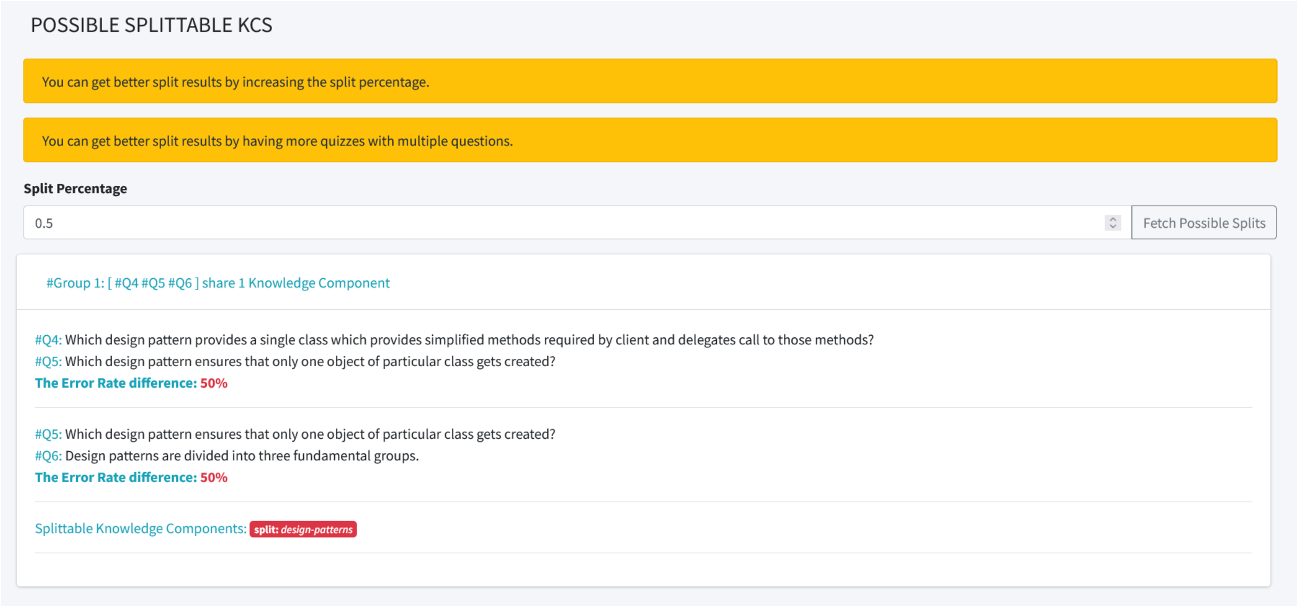


Figure IV.16: Fetch Splittable KCs Page Section

.

Manage a Knowledge Component

A teacher is able to visit a page dedicated to each KC in his course, to see all it's lessons and quiz questions.

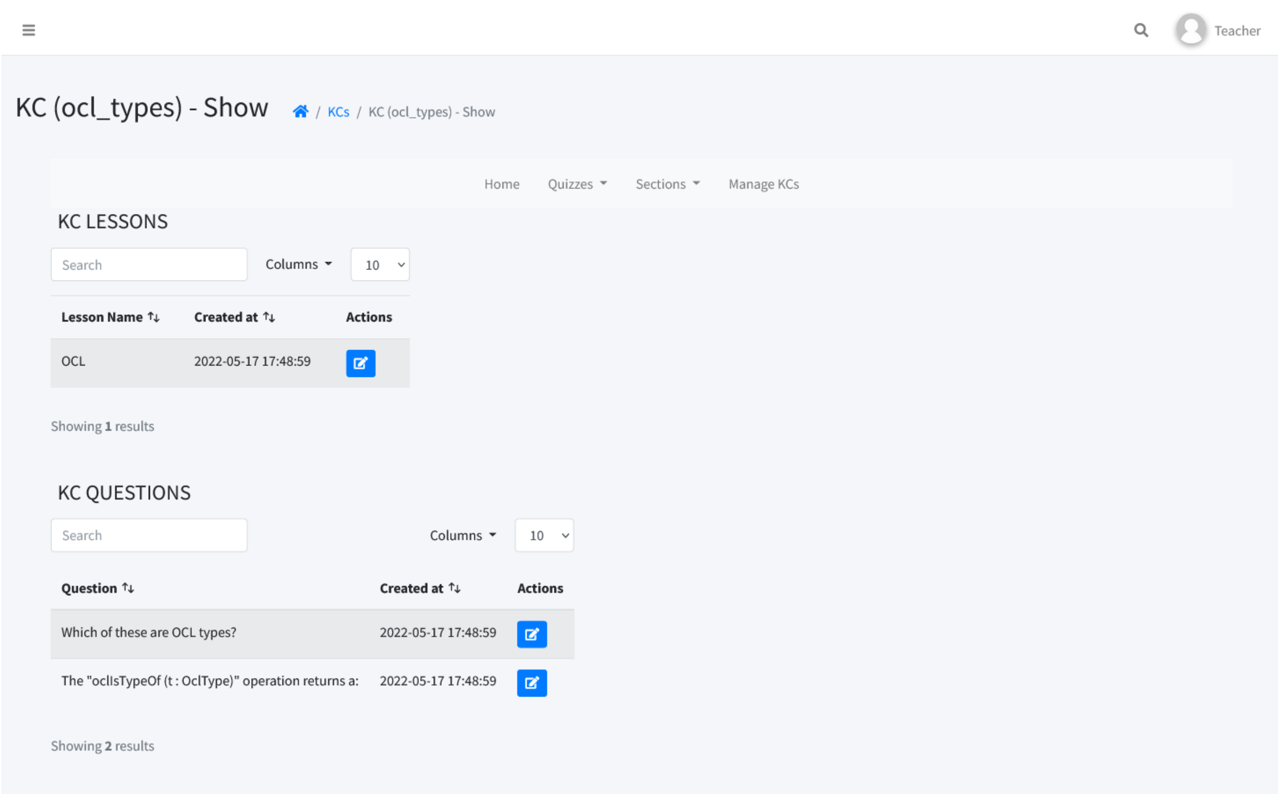


Figure IV.17: Manage One Course KC Page

.

Split a Knowledge Component

After fetching possible splittable Knowledge Components and making his choice into which one to split, a teacher can go to the split page where he will follow well defined steps to easily split a KC.

Step 1: Specify a number of splits for the chosen KC:

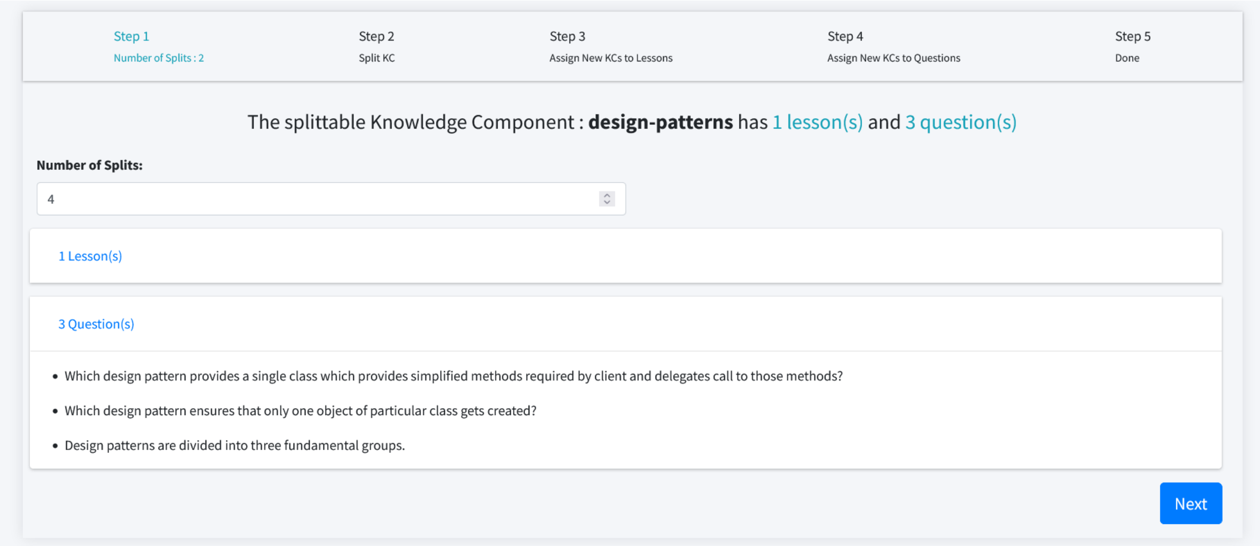


Figure IV.18: Step 1: Specify a number of splits

.

Step 2: Split the KC by creating the new replacement Knowledge Components, in our example, the teacher had previously assigned the KC 'design-patterns' to a 'Design Patterns' lesson, and 3 quiz questions related to the same subject. For this split, with the help of the dashboard, he decides to make this KC more detailed by splitting it four ways:

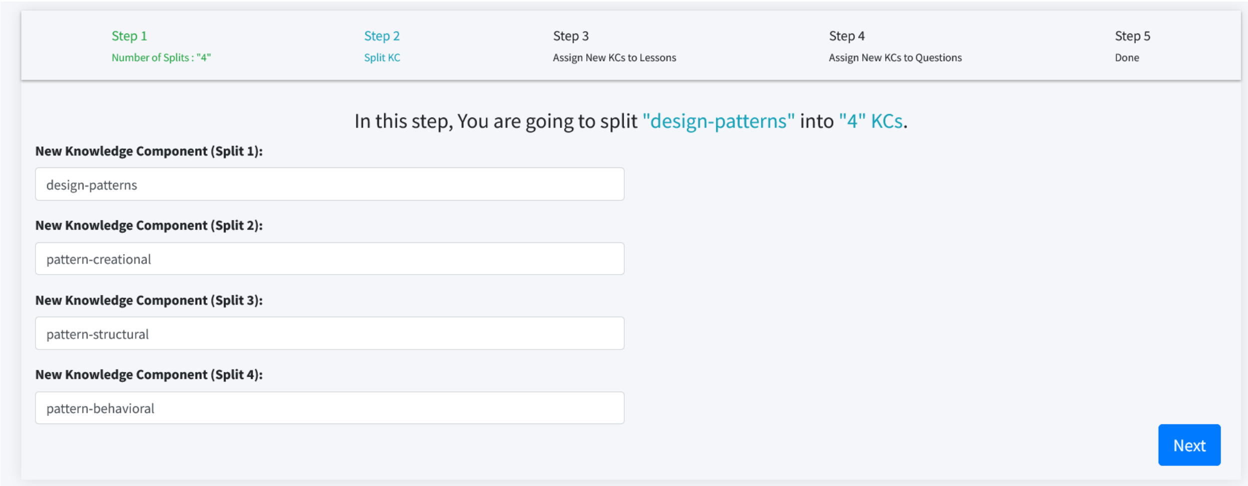


Figure IV.19: Step 2: Create New Knowledge Components

.

Step 3: Assign new KCs to old KC lessons

Figure IV.20: Step 3: Assign new KCs to old KC lessons

.

Step 4: Assign new KCs to old KC questions

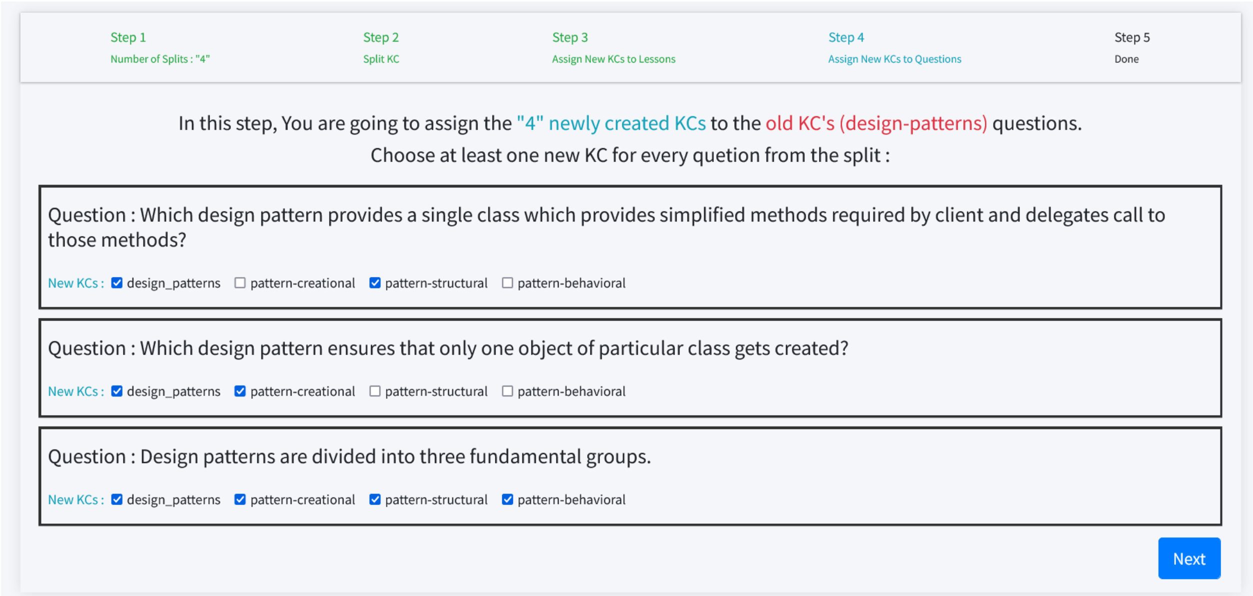


Figure IV.21: Step 4: Assign new KCs to old KC questions

.

Step 5: Validate changes

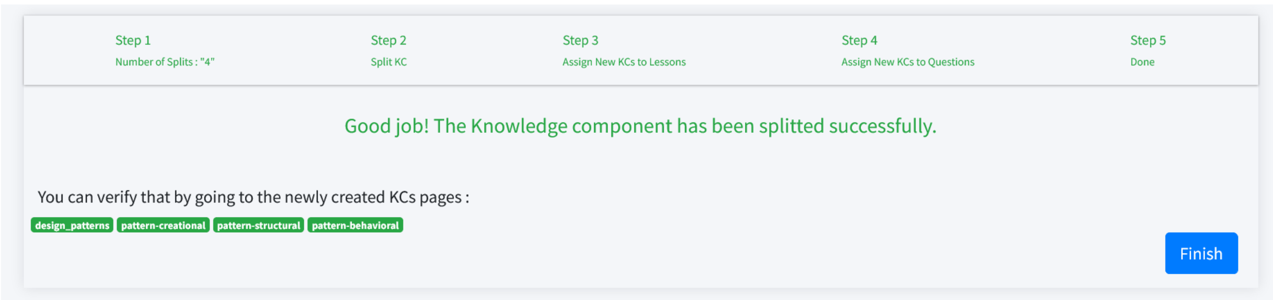


Figure IV.22: Step 5: Validate changes

.

Monitor a Quiz Page

While a student is taking a quiz, the teacher can monitor him, see his results and visualize statistics related to the quiz such as : questions' error rates, student's KC awareness (represented by three colors: green for KC fully understood. gray for either KC understanding indeterminable or unsure because the student hasn't responded to any related questions, and red for KC not understood) as shown in the figure below.

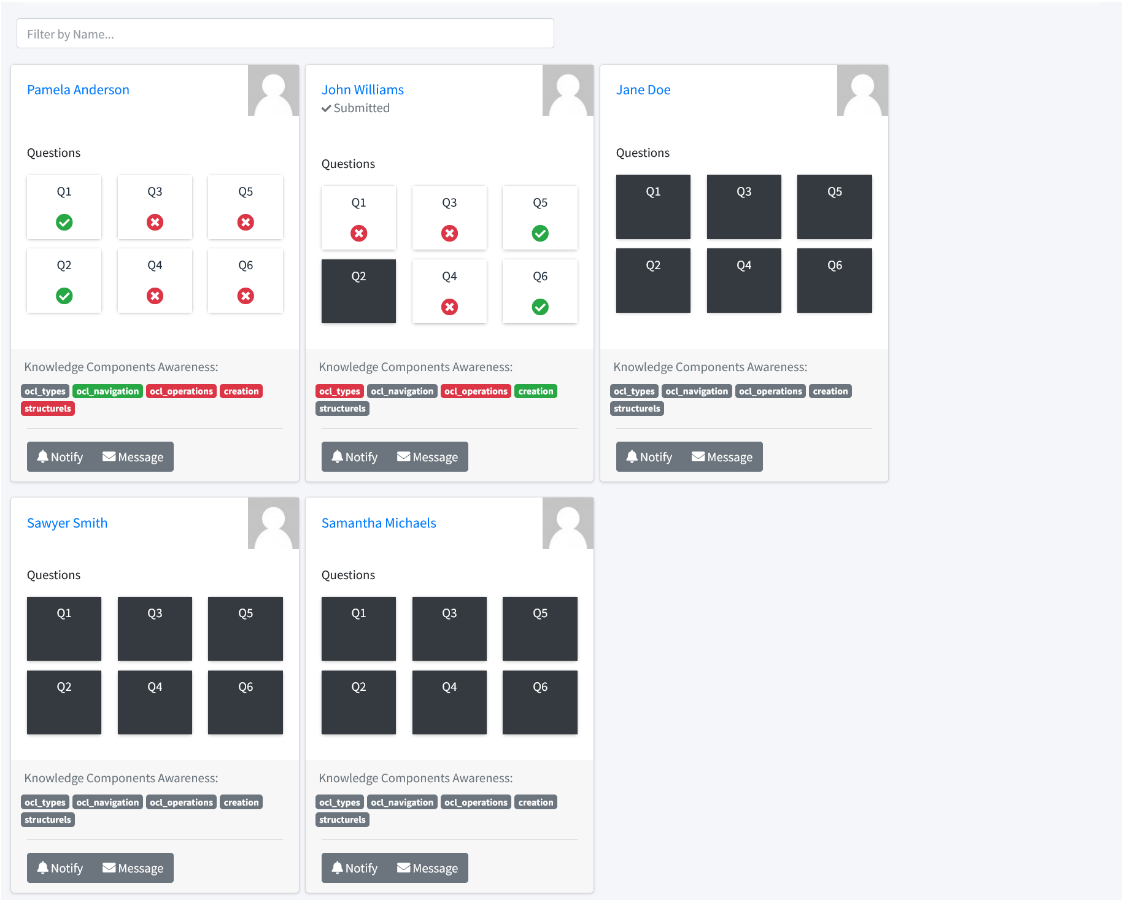


Figure IV.23: Monitor Quiz: Students View

.

Down the same page, a teacher can see two bar charts representing questions error rates, and KC error rates. In addition to that, a table with helpful real-time statistics on the quiz.

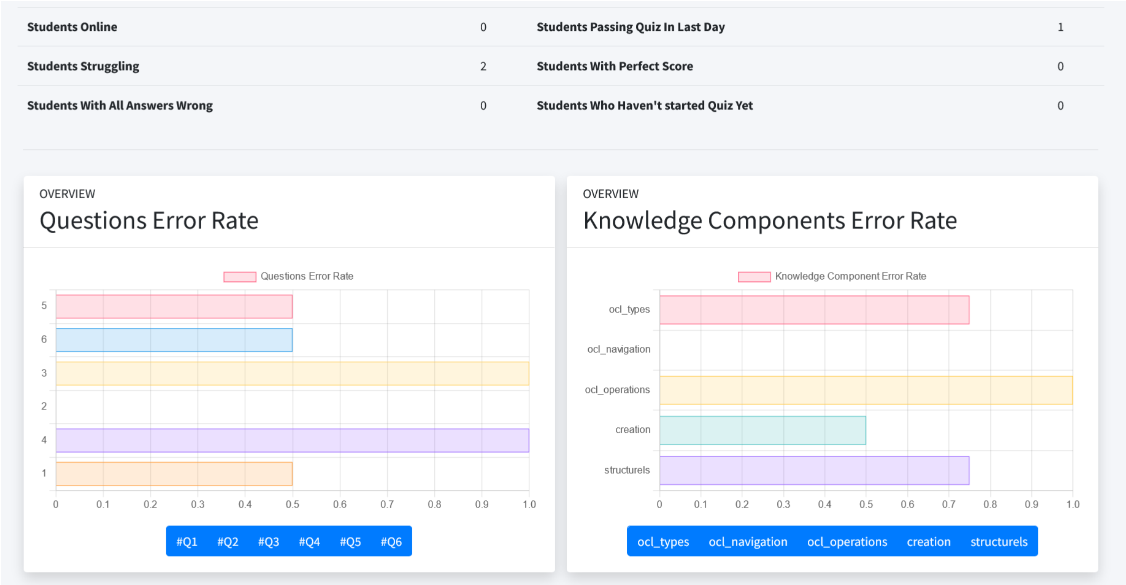


Figure IV.24: Monitor Quiz: Statistics View

.

Monitor student's KC awareness page

In this view, a teacher can see a helpful radar chart representing his student's understanding of all the knowledge components related to the quiz, in addition to a log of his responses to the quiz.

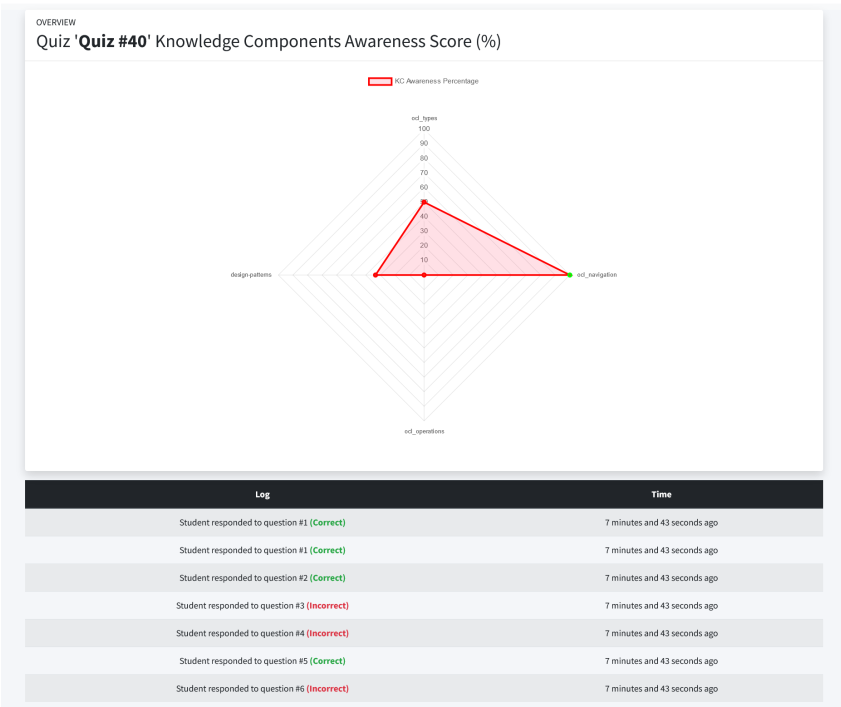


Figure IV.25: Monitor Quiz: Student's KC Awareness View

.

IV.5 Deployment

Deploying a web application can be considered the final step in the process of web development. With the help of modern tools, deploying an application and its resources to the web and securing them with Secure Sockets Layer (SSL) is an easy task.

IV.5.1 Github Actions

GitHub Actions is a tool that makes it easy to automate all software workflows, with world-class continuous integration and continuous delivery (CI/CD). Build, test, and deploy code right from GitHub. Making code reviews and managing branches are all made easy.

In our case, it was a perfect tool to make deploying to Azure web services a breeze. All that was needed was creating a YAML file in the root folder of our project that describes the GitHub Actions deployment workflow step by step.

A typical YAML github actions file is organized into three sections (figure below).

1. The *'name'* section: Gives our workflow a name
2. The *'trigger'* section: Defines a trigger for our workflow. In our case, the trigger is whenever we push our code changes to a git branch named "master".
3. The *'jobs'* section: This section of the file describes the workflow into well defined steps. It can be further divided into sections, in our case, we divided it into a "build" and "deploy" sections. Below are explanations for each step in the process:

Setup PHP: Action to set up PHP with extensions, php.ini configuration, coverage drivers, and various tools.

Check if composer.json exists: As composer.json is a critical part of every modern PHP application, checking for its existence in the root folder is a must.

Run composer install if composer.json exists: Install all PHP dependencies.

Setup Node 16.x: Install the latest long term support version of NodeJS with its package manager.

Run npm install: Installs all Node dependencies.

Run npm run prod: : Minifies all javascript and css files so that they are ready for production and are faster to read by the system.

Zip artifact for deployment: zipping the all the files before uploading them makes the process faster.

Upload artifact for deployment job: This uploads artifacts from the workflow allowing us to share data between jobs and store data once a workflow is complete.

Download artifact from build job: Downloads the shared artifact so that the workflow is continued.

unzip artifact for deployment: Unizips all files into the production environment.

delete zip file : Deleting the zip file as it is not needed anymore.

Deploy to Azure Web App: The most important step, deploying the app to Azure Web Services.

lad - master_ladapp.yml

```yaml
name: Build and deploy PHP app to Azure Web App - ladapp

on:
  push:
    branches:
      - master
  workflow_dispatch:

jobs:
  build:
    runs-on: ubuntu-latest

    steps:
      - uses: actions/checkout@v2

      - name: Setup PHP
        uses: shivammathur/setup-php@v2
        with:
          php-version: '8.0'

      - name: Check if composer.json exists
        id: check_files
        uses: andstor/file-existence-action@v1
        with:
          files: 'composer.json'

      - name: Run composer install if composer.json exists
        if: steps.check_files.outputs.files_exists == 'true'
        run: composer validate --no-check-publish && composer install --prefer-dist --no-progress --no-dev


      - name: Setup Node 16.x
        uses: actions/setup-node@v3
        with:
          node-version: 16.x
          cache: npm

      - name: Run npm install
        run: npm install

      - name: Run npm run prod
        run: npm run prod

      - name: Zip artifact for deployment
        run: zip -r release.zip ./

      - name: Upload artifact for deployment job
        uses: actions/upload-artifact@v2
        with:
          name: php-app
          path: release.zip

  deploy:
    runs-on: ubuntu-latest
    needs: build
    environment:
      name: 'Production'
      url: ${{ steps.deploy-to-webapp.outputs.webapp-url }}

    steps:
      - name: Download artifact from build job
        uses: actions/download-artifact@v2
        with:
          name: php-app

      - name: unzip artifact for deployment
        run: unzip release.zip

      - name: delete zip file
        run: rm release.zip

      - name: 'Deploy to Azure Web App'
        uses: azure/webapps-deploy@v2
        id: deploy-to-webapp
        with:
          app-name: 'ladapp'
          slot-name: 'Production'
          publish-profile: ${{ secrets.AZUREAPPSERVICE_PUBLISHPROFILE_F9DBF16520F64906805678BF0EC48A6C }}
          package: .
```

} The name of the workflow

} The action that triggers the workflow

} The steps that the workflow takes everytime it is triggered

Figure IV.26: Code: Github Actions Deployment Workflow

IV.5.2 Azure Web Services

Before deploying to any host, the most important thing to do is to see if it meets the needs of your web application. As mentioned in in the Context Diagram III.2, and explained in the "Teacher Monitors a Quiz" sequence diagram III.10, our architecture is heavily dependant on a Websockets server, to allow the real-time bi-directional communication between our back-end (the server hosting our logic) and front-end (the client, or the Learning Analytics Dashboard).

While azure web services allows this architecture to exist, it's implementation requires an experienced DevOps Engineer to execute, and our attempts to solve this problem were severely limited by our lack of understanding of what we are doing.

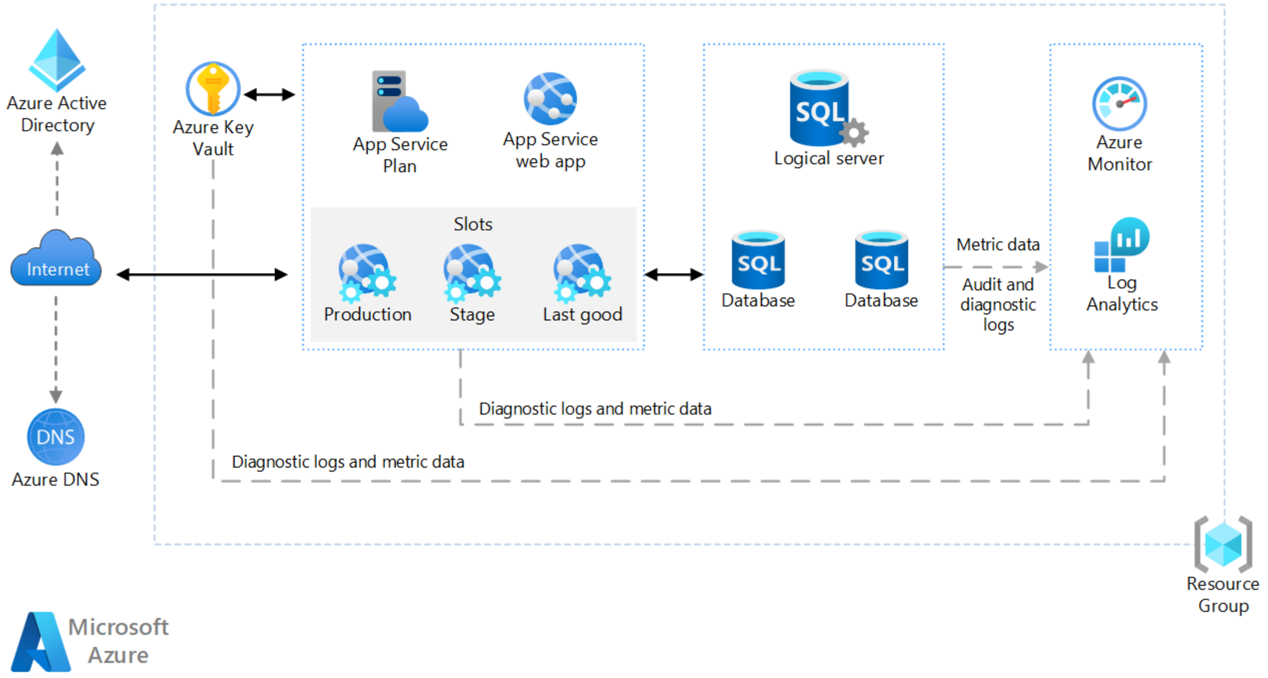


Figure IV.27: Architecture: Azure Web Service

The figure above, taken from the official Microsoft Azure Documentations[11], shows the exact architecture of our application (barring the websockets server), every resource in this service is grouped inside a "Resource group", which is Azure's way to know in which region to put your resources, and how to bill you for every transaction coming into your resources.

Then, our resource group was split into 3 parts, the "app service web app" which contains our actual web application, as shown in the figure above, it may contain multiple deployment slots that let you stage a deployment before swapping it with the production deployment slot.

Then comes the Azure SQL Database, which is Database-as-a-Service that is well integrated into Azure web

[11] https://docs.microsoft.com/en-us/azure/architecture/reference-architectures/app-service-web-app/basic-web-app?tabs=cli

services. In there, hosting our data, and student activity logs was done effortlessly thanks to the Azure Portal[12] dashboard, which simplified managing our resources.

Finally came the Log Analytics and Azure Monitor services, that come automatically enabled within every Azure Web Service, and are a great way for the website admin to have a better understanding of several metrics associated with the performance of every web app, and can be even scoped to include analysis on every other part of the resource group.

IV.6 Conclusion

In this chapter, we have presented the implementation of the system we have developed as well as a global description of our application. and we have presented the result using the interfaces. Although we are satisfied with our progress, we think that we need to do more global tests using a much larger amount of data. However, the implementation is not going to stop here, we plan to continue adding new features and improving existing ones.

[12]https://portal.azure.com/

General conclusion

In this project, we were interested in creating an application of Learning Analytics Dashboard (LAD) in Higher Education (HE) to monitor students' learning in real-time. This particularity will allow teachers to better understand the students' progress, optimize the learning experience and improve teachers' relations and students.

Firstly, we started by defining Learning Analytics and studying its history . We've made sure to mention its different steps according to some of the latest research, shed light on each of them and list metrics Used in that field. Then a deep dive was made in the theoretical notions used in LA , where an introduction machine learning and data mining has been made.

Secondly, we reviewed recent research studies into learning data analytics, LA and EDM techniques and trends in building predictive or explanatory models that use data and also the various solutions to improve the learning process in Higher Education. Then, we have pointed out LA challenges and opportunities.

Which brings us to the third chapter of this research, where we have pointed up the problems to solve after that our potential solution and the different design steps of the Real-time Learning Analytics Dashboard using the Knowledge Components Algorithm. Then, we proceeded to a detailed conceptual study of our application based on a variety of UML diagrams.

Then, in the fourth chapter we went on the implementation phase where we applied the theoretic notions and methods explored in past chapters to build our LAD taking great care to choose robust technologies to ensure an effective application. It is important to note that our work remains perfectible. Indeed, some points that we have not been able to implement in spite of time, can be improved in the future to have an application as operational as possible. In the following we present the perspectives of this project. In the short term :

- Prevent identity theft attempts by sending a validation email to the email account linked to the user's account.
- Send notifications and reminders to users' main emails, a feature they can enable or disable from the user's

In the long term:

- To reach the level of operationality we are looking for, many tests are necessary, and these tests are needed, and this using a very large amount of data. Perhaps the way to achieve this is through beta testing, where a number of students and teachers are invited to use the application for a year to simulate the process and send bug reports and feedback throughout their use, which will be corrected by the application's maintainers.

Bibliography

[1] Elias, T. (2011). Learning analytics: Definitions, processes and potential (Report). Retrieved from http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf

[2] Leitner, Philipp Khalil, Mohammad Ebner, Martin. (2017). Learning Analytics in Higher Education—A Literature Review. 10.1007/978-3-319-52977-6_1

[3] O. Skrypnyk, JoksimovicS., KovanovicV., GasevicD. and S. Dawson. (2015). Roles of Course Facilitators, Learners, and Technology in the Flow of Information of a CMOOC. International Review of Research in Open and Distance Learning. 16. 188. 10.19173/irrodl.v16i3.2170.

[4] Joksimovic, Srecko Kovanovic, Vitomir Dawson, Shane. (2019). The Journey of Learning Analytics. 6. 37-63.

[5] Campbell, J.P., DeBlois, P.B. Oblinger, D.G. (2007). Academic Analytics: A New Tool for a New Era. EDUCAUSE Review, 42(4), 40-42. Retrieved March 1, 2022 from https://www.learntechlib.org/p/100331/.

[6] Paiva, Ranilson Bittencourt, Ig Lemos, Wansel Vinicius, André Dermeval, Diego. (2018). Visualizing Learning Analytics and Educational Data Mining Outputs. 10.1007/978-3-319-93846-2_46.

[7] Narciss, Susanne Huth, Katja Narciss, Dr. (2002). How to design informative tutoring feedback for multimedia learning. Instructional Design for Multimedia Learning.

[8] H. Aldowaha, H. Al-Samarraiea and W. M. Fauzy, "Educational data mining and learning analytics for 21st century higher education: A review and synthesis," Telematics and Informatics, vol. 37, pp. 13-49, 2019. 10.1016/j.tele.2019.01.007.

[9] Arnold, Kimberly Pistilli, Matthew. (2012). Course signals at Purdue: Using learning analytics to increase student success. ACM International Conference Proceeding Series. 10.1145/2330601.2330666.

[10] Saqr, Mohammed. (2018). Using Learning Analytics to Understand and Support Collaborative Learning. 10.13140/RG.2.2.29579.44321.

[11] J. Kuzilek, M. Hlosta, D. Herrmannova, Z. Zdrahal and J. a. W. A. Vaclavek, "OU Analyse: Analysing at-risk students at the Open University.," in International Learning Analytics and Knowledge (LAK) Conference, New York, 2015.

[12] C. Brooks and C. Thompson, "Predictive Modelling in Teaching and Learning," in Handbook of Learning Analytics, Michigan, SOLAR Society for Learning Analytics Research, 2017, pp. 61-68.

[13] R. Liu and K. R. Koedinger, "Going Beyond Better Data Prediction to Create Explanatory Models of Educational Data," School of Computer Science, Carnegie Mellon University, USA, Pennsylvania, 2017.

[14] "DataShop Cheat Sheet," DataShop, July 2014. [Online]. Available: https://pslcdatashop.web.cmu.edu/downloads/DataShopCheatSheet.pdf. [Accessed 14 February 2022].

[15] Chatti M.A., Muslim A., Guliani M., Guesmi M. (2020) The LAVA Model: Learning Analytics Meets Visual Analytics. In: Ifenthaler D., Gibson D. (eds) Adoption of Data Analytics in Higher Education Learning and Teaching. Advances in Analytics for Learning and Teaching. Springer, Cham. https://doi.org/10.1007/978-3-030-47392-1_5

[16] Keim D., Andrienko G., Fekete JD., Görg C., Kohlhammer J., Melançon G. (2008) Visual Analytics: Definition, Process, and Challenges. In: Kerren A., Stasko J.T., Fekete JD., North C. (eds) Information Visualization. Lecture Notes in Computer Science, vol 4950. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-70956-5_7

[17] Gibson, D. Ifenthaler and David, "Adoption of Learning Analytics," in Adoption of Data Analytics in Higher Education Learning and Teaching, Mannheim, Springer, 2020, pp. 3-16.

[18] Gasevic, Dragan Joksimovic, Srecko Eagan, Brendan Shaffer, David. (2018). SENS: Network analytics to combine social and cognitive perspectives of collaborative learning. Computers in Human Behavior. 92. 10.1016/j.chb.2018.07.003.

[19] Arthars, Natasha Dollinger, Mollie Vigentini, Lorenzo Kondo, Elsuida King, Deborah. (2018). Empowering teachers to personalize learning support. 10.31234/osf.io/3xu5t.

[20] Blackmon S.J., Moore R.L. (2020) A Framework to Support Interdisciplinary Engagement with Learning Analytics. In: Ifenthaler D., Gibson D. (eds) Adoption of Data Analytics in Higher Education Learning and Teaching. Advances in Analytics for Learning and Teaching. Springer, Cham. https://doi.org/10.1007/978-3-030-47392-1_3

[21] A. F. Wise and J. Vytasek, "Learning Analytics Implementation Design," in Handbook of Learning Analytics, Michigan, SOLAR Society for Learning Analytics Research, 2017, pp. 151-160.

[22] Anderson, C. (2008, 23 June). The end of theory: The data deluge makes the scientific method obsolete. Wired Magazine. http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory