

Andine Luick & Fredrik Monsen

Visualizing Data from Adaptive Self-Assessment to Support Teaching and Learning Activities

Master's thesis in Master of Science in Informatics

Supervisor: Michail Giannakos

Co-supervisor: Zacharoula Papamitsiou

June 2022

Andine Luick & Fredrik Monsen

Visualizing Data from Adaptive Self-Assessment to Support Teaching and Learning Activities

Master's thesis in Master of Science in Informatics

Supervisor: Michail Giannakos

Co-supervisor: Zacharoula Papamitsiou

June 2022

Norwegian University of Science and Technology

Faculty of Information Technology and Electrical Engineering

Department of Computer Science



Norwegian University of
Science and Technology

Abstract

An emerging tool used by both learners and tutors is Learning Analytics Dashboards (LADs). Students' learning data can be collected, aggregated, and visualized to provide both learners and tutors alike with actionable information for assessment, intervention, and adaptation. Previous studies have reported mixed results for such visualizations, with teachers struggling to interpret the presented information, although if this information is displayed in a concise and actionable format, it has been found that these systems help teachers identify students at risk and monitor progress and learning.

In this thesis we have reviewed the literature to collect the requirements of a teacher-facing dashboard, extending the pre-existing SmartU student dashboard. Using the design and creation strategy we have developed a web application. This became the artifact of a user test where participants with teaching experience used the system under observation. Data was also collected from interviews and a questionnaire. Our hypothesis is that an LAD can effectively support teachers' objectives. The result of the study indicates that LADs can support teachers in identifying struggling students or issues in the learning material. However, differences were found between teachers and implies dashboards might not fit with every type of tutoring, or that the dashboard must be customizable to fit with different types of education.

Sammendrag

Dashbord for læringsanalyse er en fremvoksende type verktøy som kan brukes av både lærere og elever. Læringsdata om studenter kan samles inn, aggregeres og visualiseres for å gi både elever og lærere praktisk informasjon for vurdering, intervensjon og tilpasning. Tidligere studier har rapportert om blandede resultater for slike visualiseringer, med lærere som sliter med å tolke informasjonen som er presentert. Dersom denne informasjonen vises i et kortfattet format som legger til rette for handling så er det funn som tyder på at slike systemer kan hjelpe lærere i å identifisere studenter og elever som er i faresonen for å stryke, samt overvåke progresjon og læring.

I denne oppgaven har vi gjennomgått litteraturen for å oppsummere kravene til et dashbord for læringsanalyse for lærere, og utvide det eksisterende SmartU elev- og studentdashbordet med et lærerdashbord. Ved å bruke en design og utviklingsstrategi har vi utviklet en webapplikasjon for lærere. Denne webapplikasjonen ble testet gjennom en brukertest der deltakere med undervisningserfaring brukte systemet under observasjon. Data ble også samlet inn fra intervjuer og spørreskjema. Vår hypotese er at et dashbord for læringsanalyse effektivt kan støtte lærernes mål, og studiens resultater indikerer at slike dashbord kan støtte lærere i å identifisere problemer med lærestoffet og elever som sliter. Imidlertid ble det funnet forskjeller mellom lærere som antyder at dashbordet kanskje ikke passer med alle typer undervisning, eller at dashbordet må kunne tilpasses for å passe med ulike typer undervisningsformer og utdanningsnivå.

Acknowledgment

This master thesis was written between fall 2021 to spring 2022 written for the Department of Computer Science (IDI) at the Norwegian University of Science and Technology (NTNU).

We would like to give a special thanks to our supervisors Michail Giannakos and Zacharoula Papamitsiou for their continued feedback, guidance and encouragement throughout the project.

We would also like to thank our families and friends for their support. Without their support we could not have succeeded in completing this project.

At last but not least, we would also like to direct thanks to all those who participated in the user study, which made this thesis possible.

Andine Luick & Fredrik Monsen
Trondheim, June 10, 2022

Table of Contents

List of Figures	viii
List of Tables	ix
1 Introduction	1
1.1 Problem description	1
1.2 Goals and research questions	2
1.3 Research design & research methods	2
1.3.1 Research strategies	2
1.3.2 Data generation methods	3
1.3.3 Data analysis	3
1.4 Contributions	3
1.5 Thesis structure	3
2 Background theory and related work	5
2.1 Literature review	5
2.2 Previous studies	6
2.2.1 Features	6
2.2.2 Adapting learning format	7
2.2.3 Intervention	8
2.2.4 Visualization	9
3 Design and implementation	11
3.1 Functional requirements	11

3.2	Tools	13
3.2.1	Design tools	13
3.2.2	Development tools	13
3.3	Design process	13
3.3.1	First iteration	14
3.3.2	Second iteration	15
3.3.3	Figma prototype	17
3.3.4	Final design	19
3.4	Current system	27
3.4.1	System architecture	27
4	Methods	29
4.1	Study design	29
4.1.1	Participants	29
4.1.2	Test setup	30
4.1.3	Test procedure	31
4.2	Data collection	32
4.2.1	Observation	32
4.2.2	Questionnaire	32
4.2.3	Interviews	33
4.3	Data analysis	33
4.3.1	Qualitative data analysis	34
4.3.2	Quantitative data analysis	37
5	Results	40
5.1	Results for qualitative data analysis	40
5.1.1	Visualizations	40
5.1.2	Intervention	44
5.1.3	Perceptions towards notifications and highlighting	48

5.1.4	Content management	50
5.1.5	Privacy	51
5.1.6	Customization	51
5.1.7	Attitude towards using the LAD	52
5.2	Results for quantitative data analysis	54
5.2.1	Testing for assumptions	54
5.2.2	Descriptives	54
5.2.3	Correlation	55
5.2.4	Mann-Whitney U test	56
5.2.5	Usability	56
6	Discussion	57
6.1	Effectively supporting teacher intervention	57
6.2	Effectively supporting teachers in adapting their lesson plan and learning format	58
6.2.1	Perceptions towards the different visualizations	58
6.2.2	Perceptions towards content management	61
6.2.3	Perceptions towards customization	61
6.3	Attitude towards an LAD	62
6.4	Implications	62
6.5	Limitations	63
7	Conclusions and future work	64
7.1	Future work	65
	Bibliography	66
	Appendices	69
	A Study	70
A.1	Form of consent	71

A.2	Observation guide	74
A.3	Interview guide	75
A.4	Questionnaire	76
B	Interviews	86
C	Design and implementation	121
C.1	User tests	121
C.2	Final user interface	128
C.3	Code repository	141

List of Figures

1	Iteration 1 virtual pen & paper design	15
2	Iteration 2 activity dashboard	16
3	Iteration 2 student details	17
4	Figma prototype of dashboard	18
5	Figma prototype of landing page	19
6	Activity dashboard	20
7	Application landing page	21
8	Question management and overview	22
9	Edit activity	23
10	Edit question	24
11	Student views	25
12	Application flow	26
13	System architecture	27
14	Demographics	30
15	The interviewees' teaching level	30

List of Tables

1	Overview of topics and related papers	6
2	Functional requirements	12
3	Age descriptives	30
4	Observation guide	32
5	The initial 8 categories	36
6	The final 7 categories with subcategories	36
7	Constructs	39
8	General thoughts about the dashboard design	40
9	Statements about the highlighted statistics	41
10	Statements about the learning progress graph	42
11	Statements about the Overview visualizations	43
12	Statements about the Mastery Grid	43
13	Statements about contacting students	44
14	Statements about changing lecture plan, curriculum etc.	46
15	Statements about the notifications	48
16	Statements about content management	50
17	Statements about privacy	51
18	Statements about customization	51
19	Statements about how the student scores should be sorted	52
20	Attitude towards using the LAD	52
21	Descriptives for PEOU and PU items	55
22	Spearman correlation matrix	55
23	Mann-Whitney U test results	56

24	Result of SUS items	56
----	-------------------------------	----

Acronyms

LA Learning Analytics. 1, 5, 8, 9, 37

LAD Learning Analytics Dashboard. 1–8, 35, 36, 44, 45, 47, 50, 53, 57, 58, 61, 64

PEOU Perceived Ease of Use. 3, 32, 33, 37, 54

PU Perceived Usefulness. 3, 32, 33, 37, 54

SmartU Self-assessment Measured with Analytics on Run-Time for YOU. 1, 2, 4, 13, 19, 21, 58, 59, 61, 62, 64, 65

SUS System Usability Scale. 3, 32, 37, 56, 59

1 Introduction

This thesis describes the research, development, and testing of the teacher-facing Learning Analytics Dashboard (LAD) for the self-assessment application SmartU. Designing, developing, and evaluating a learning analytics dashboard can be a difficult task. This section explains the problem and rationale for performing the research, the objectives and research questions, the methods utilized to achieve these objectives, and the research's contributions. Finally, an overview of how the thesis is organized is presented.

1.1 Problem description

SmartU (Self-assessment Measured with Analytics on Run-Time for YOU) is an adaptive self-assessment system for students using Learning Analytics (LA). In SmartU students can take quizzes which adapts to their skill level, earn achievements and measure score in comparison with their peers [1]. The system does not contain a separate view for teachers, who could benefit from displaying the data generated by student interaction with the system.

LA concerns itself with “*the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs*” as defined during the first Learning Analytics and Knowledge conference in 2011 [2]. Dashboards using LA, referred to as Learning Analytics Dashboards (LADs), are user interfaces which present information in form of graphs, charts, and other visualization techniques to end-users such as teachers and students [3]. The aim is to support both learning and teaching objectives, such as supporting struggling students, assessing the class, or adapting parts of the curriculum.

The data of students' learning collected from predictive LA systems can provide teachers with essential information for assessment and intervention [4]. Previous studies have reported mixed results from visualizations of the data from such systems to teachers [5]. Some of the previous studies found that most teachers struggled to understand the visualizations and results presented from the various systems [5]. On the other hand, if information is presented in a concise and actionable format, LADs can help teachers to identify students at risks as well as monitor the students' progress and learning [4, 6]. This can have an effect on the teachers' decision to alter their lecture plan according to the students' mastery level [6].

The hypothesis for this thesis is that an LAD can effectively support teacher intervention and adaptation of learning format. The thesis will investigate the hypothesis with

its research questions which are presented in the following subsection. A study will be conducted to get insights into how teachers perceive an LAD, and if the highlighted data actually help them to identify important and critical data or if it is just distracting and a nuisance.

1.2 Goals and research questions

To verify the hypothesis, this thesis attempt to answer the following research questions:

- **RQ1:** *How can an LAD effectively support teacher intervention?*
- **RQ2:** *How can an LAD effectively support teachers in adapting their teaching, lesson plan and learning format?*
- **RQ3:** *What are teachers' motivation and attitudes towards using an LAD?*

1.3 Research design & research methods

In order to answer the research questions, the research was planned according to Oates' methods [7], selecting a research strategy, data generation methods and planning the data analysis. To provide the conceptual framework for the research, a literature review was conducted. The conceptual framework helps the researchers to identify factors that comprise the topic and the way of thinking about the topic. It also makes it clear what strategies and methods and data analysis to use and how to design and create the LAD as well as how to evaluate the research.

1.3.1 Research strategies

Since the thesis consists of both developing an LAD and testing it, the *Design and creation* strategy according to Oates [7] was used combined with user tests with observations. To design and develop the software artifact, the *Design and creation* strategy was applied. Oates [7] describes a five-step plan that consists of the awareness, suggestion, development, evaluation and conclusion. These steps are not strictly followed but applied in a way that results in gathering insight through literature, designing the software artifact in iterations, and lastly testing and evaluating the developed artifact. Details about the steps of designing the artifact can be seen in Section 3.

To test the SmartU LAD and identify the system's usefulness and ease of use, user tests were performed. The user tests were conducted in a controlled environment to avoid variables, in a similar manner to Oates' suggestion for the Experiment strategy. The user tests were only conducted once with no control group, and therefore the Experiment strategy could not be applied.

1.3.2 Data generation methods

Data generation methods are often associated with particular research strategies. By using more than one method, more data will be produced and potentially improve the quality of the research [7]. In addition, findings from one data generation method can be used to corroborate by comparing it with data from another method. Because of this, the data generation methods used for this thesis is interviews, observations, and questionnaire. The use of multiple data generation methods is also referred to as method triangulation. The triangulation was used to cross-check the results by validating the findings or finding contradictions.

1.3.3 Data analysis

Since both qualitative and quantitative data was collected, approaches for qualitative and quantitative data analysis were applied. The transcribed interviews were coded into categories using Mayring's model of inductive category development [8] and observations were compared to the interview findings for cross-checking. The data collected from the questionnaire was used to calculate the System Usability Scale (SUS) by using Brooke's method [9]. The frequency, minimum, maximum, mean and standard deviation were calculated for each of the Perceived usefulness (PU) and Perceived ease of use (PEOU) constructs. A Mann-Whitney U was also carried out.

1.4 Contributions

This thesis contributes with a design and creation strategy including user tests, evaluated with a combination of qualitative and quantitative empirical results. The research collects the findings of previous studies to create a new dashboard as an artifact for a user test, to evaluate the combination of features and visualizations found in the literature. The results highlight the strengths and weaknesses to the teacher-facing LAD, by providing insights into how participants perceived the different aspects of the system, such as the different visualizations and how it can help them achieve their objectives as teachers. These insights point out what features that are helpful and features that potentially can be improved or added, which leads to discussions about future work.

1.5 Thesis structure

- **Section 1:** Introduction – presents the problem description, the research goals and questions, the research methods used to address the questions, and the contributions of this thesis.
- **Section 2:** Background theory and related work – provides definitions of relevant terms, describes the steps and findings of the literature review and as well as information about features used in previous teacher-facing LADs based on literature.

- **Section 3:** Design and implementation – describes the iterative design and the agile development processes followed for creating the teacher-facing LAD for the SmartU system.
- **Section 4:** Methods – describes the methods used in the study: The design of the study, the data collection and data analysis methods.
- **Section 5:** Results – presents the results from the qualitative and quantitative data analysis.
- **Section 6:** Discussion – discusses the findings and compares the results of this study to previous and related studies.
- **Section 7:** Conclusions and future work – presents the conclusions of this thesis.

2 Background theory and related work

2.1 Literature review

Prior to the literature search, we familiarized ourselves with the topic of LA and LADs using articles provided to us by the supervisors, to generate relevant search terms. We decided to search for articles about LADs, specifically their implementation and usage by teachers, as well as articles discussing adaptive self-assessment and adaptive tutoring dashboards. The literature review does not differentiate between studies from various education levels.

The motivation for the literature review is to identify requirements of a teacher-facing dashboard, such as essential features, what to visualize, and how it is being used by teachers. Another important motivation is to identify issues with teacher dashboards in the existing literature.

To conduct the search, Google Scholar and ACM Digital Library were used. While Google Scholar gives a wide array of results, ACM contains conference proceedings from the LAK conferences, which contains high quality research within the specific field of LA.

Search terms: *“learning analytics dashboard” / “learning analytics dashboard” AND (teachers OR tutors OR lecturer) / “learning analytics” AND (teachers OR tutors OR lecturer) / “adaptive self-assessment” AND (teachers OR tutors OR lecturer) / “adaptive tutoring dashboard”*

The above search terms were used to find articles which specifically focused on LADs and systems using LA from a more teacher-centric perspective, as searching without appending *“teachers”* gave more general articles which often focused on learners. The initial search resulted in 42 papers.

To reduce the initial list of papers we set up some criteria. One of the following criteria had to be met:

1. Focuses on a teacher’s dashboard
2. Focuses on LA from a teacher’s perspective

Papers that did not fit either of these criteria were discarded. Furthermore, the remaining papers were scanned to identify common topics: features, adaption of the learning format, intervention, and visualization. Papers that did not cover any of these topics sufficiently were also discarded. This reduced the total papers used down to 16.

The results of the literature review are presented in the categories of features, adapting the learning format, intervention, and visualization. Table 1 illustrates the identified papers per topic.

Features	Herodotou et al. [10], Herodotou et al. [5], Isaias and Viana [11], Michos et al. [12], Heffernan et al. [13], Molenaar and Knoop-van Campen [14]
Adapting learning format	Khakaj et al. [6], Michos et al. [12], Heffernan et al. [13], Alevan et al. [15], Khakaj et al. [16],
Intervention	Herodotou et al. [10], Khakaj et al. [6], Molenaar and Knoop-van Campen [14], De Freitas et al. [17], Shibani et al. [18]
Visualization	Herodotou et al. [10], Zheng et al. [4], Khakaj et al. [6], Isaias and Viana [11], Michos et al. [12], Feng et al. [19], Bao et al. [20], Dourado et al. [21], Park and Jo [22]

Table 1: Overview of topics and related papers

2.2 Previous studies

2.2.1 Features

Based on habits and experiences, the way we perceive different features in an LAD can be subjective. Notifications for critical information has been a missing dashboard feature that most teachers agree on in previous studies [10, 5, 11, 12, 14]. Herodotou et al. [5] and Isaias and Viana [11] found from interviews with teachers that they find it stressful and time-consuming to login and manually check for students at risk regularly. Interviews conducted by Michos et al. [12] found that teachers believed automatically generated reminders as well as sending updates to students regarding assignments or participation can help improve their workflow and efficiency. Isaias and Viana [11] also found that the teachers rather would want the system to automatically send reminders to the students of risk instead of this being an extra responsibility for them. This is the opposite of Herodotou et al.’s [10] findings in another study, where the majority of the teachers found this kind of information helpful for reaching out to and assisting the students in need.

In addition to only adding notifications, Molenaar and Knoop-van Campen [14] suggested improving LADs by adding new services and highlighting important dashboard data to make the pedagogical actions more explicit using recommender services. This was due to instructors not knowing what to do with the information displayed on the LAD. By adding recommender services, it can give the instructor suggestions for what actions to take according to the information received.

Functionality for adding, changing, and removing content can be useful for a teacher’s dashboard for an adaptive learning system. This is supported by a study conducted by Heffernan and Heffernan [13], addressing the history and 8-year usage of a learning platform called ASSISTments. It was found that teachers find it useful to have the possibility

to add and change/replace content, such as adding questions to a multiple-choice quiz or adding hints to a question.

2.2.2 Adapting learning format

LADs can provide teachers with feedback about their students' progress, helping them reflect and make decisions regarding lectures and learning format [23]. Alevan et al. [15] made a list of analytics that can be useful for adaptive tutoring, based on knowledge from previous studies and teacher reports. The analytics elements were divided into four groups; (1) progress (overall progress such as a list of completed activities or progress during a specific session), (2) skill mastery and learning rate, (3) learning behaviors (effort, time spent, use of hints or other gamification elements) and (4) challenges (which problems are hardest etc.).

When designing a teacher's dashboard for adapting lectures, it is interesting to look at what data motivates the teacher to change the lecture plan. Xhakaj et al. [16] wanted to inspect how different information displayed on the intelligent tutoring system Luna can affect the teacher's tutoring plan. The available information on the dashboard was (1) the number of students who had mastered each skill, (2) the number of students who made certain errors, and (3) a scatter plot with a comparison of the mastery level versus the amount of practice per skill, averaged across students. In a study of 8 experimental condition classes using a high-fidelity prototype of Luna, they collected information that showed that the dashboard affected the teacher's knowledge about the class overall as well as individual students [16]. Results from the experiment also showed that the dashboard made the teachers change their lesson plans, mostly because of the struggling students. In 1/8 classes, the teacher also decided not to cover a topic because the class had already mastered it. Michos et al. [12] initiated two projects that aimed to identify which data the teachers believed could be useful, where their skills are coming up short, and where technology is needed. The results for the projects showed that they wanted functionality like identifying areas of student success on tasks and predicting difficulty on classroom activities.

Adapting the learning format according to the strengths and weaknesses of the class requires the teacher to repeatedly define issues. Identifying misconceptions can be just as useful as identifying mastered concepts for adapting the lecture plan. This is supported by another study by Xhakaj et al. [6], where contextual inquiry interviews were conducted on 6 teachers from 3 different schools to study how teachers use student data to adapt lectures or individualize instructions. The findings from the interviews revealed that the teachers regularly generate and use data on their students' progress and performance. The overall class or individual mastery of concepts is something the teacher especially pay attention to as well as misconceptions for the different assignments [6]. Heffernan and Heffernan [13] had similar findings, as they found information about how many students answered a question correctly or wrongly useful to identify misconceptions for teachers to address or use as a starting point for a meaningful classroom discussion. Michos et al. [12] found in their study with LADs that the teachers experienced difficulty identifying what the class

was struggling with. Another challenge was dividing the issues into smaller parts to be able to solve them.

2.2.3 Intervention

An important usage area of LADs is for teachers to identify students at risk of failing. Intervention can also help retention by having teachers assist students at an early stage when problems start to arise. Predictive LA can be used to identify these students. OU Analyse is a predictive LA dashboard which uses a machine learning algorithm to identify students at risk, making use of static data such as demographics, education, and previous course scores, as well as static data which consist of interactions with the system [10]. Herodotou et al. [10] found that only weekly usage and best previous course score of students were significant in their prediction in the predictive LA system. Xhakaj et al. [6] found through their contextual inquiry interviews that collected student data is used to make class-level and individual-level decisions with regards to intervention. As mentioned earlier, the teacher can intervene by adapting the lectures to spend more time with difficult concepts that several students have failed to master, and can target individuals that are found to be struggling. Shibani et al. [18] found that in addition to helping the teachers give feedback to the students and intervene when necessary, using LA for intervention became disrupting and time consuming.

To identify students at risk of failing, it is crucial to use the correct data to power the output. De Freitas et al. [17] found several data that can be used to power retention related prediction when they tested five hypotheses about retention and attrition. The hypotheses were discovered from a dataset of 51 181 students at an Australian university. They found support for three of the five hypotheses: (1) engaging in online lecture materials while also having on-campus attendance reduced chance of attrition, (2) satisfaction with one's own academic performance reduced attrition, and (3) being closer to the average age of one's cohort increased chance of retention. While not an exhaustive list, it gives indication of some of the data that may be required to power accurate predictions.

With a study observing teachers in primary schools, Molenaar and Knoop-van Campen [14] describe the implementation of a learning analytics dashboard in practice. Students in the study used a tablet with an adaptive educational system that recorded performance data in real time. Teachers either did nothing, gave feedback, instruction, or altered learning materials or pace after interacting with the dashboard. These activities were labeled as pedagogical. Following consultation of the dashboard, 50% of pedagogical interventions were targeted at individual students, 7% at small groups, and 43% at the class level. Individual students were given comments and directions most of the time, but class activities were mostly instructions. A surprising conclusion was that a quarter of dashboard consultations resulted in no action from the teacher. According to the authors, this could be due to teachers reinforcing previous impressions of student and class success [14].

2.2.4 Visualization

There is some uncertainty regarding what aspects are more suitable for the teacher dashboard to ensure a clear understanding of how learning occurs. User-friendliness of the visual representation can lead to a better interpretation of the collected data [11]. It is important that the dashboard displays LA that includes different types of metrics. From semi-structured interviews conducted by Isaias and Viana [11], 14 teachers were asked about what information they thought would be useful for the dashboard. The results from the interviews were organized in three categories: (1) descriptive data, (2) comparative data, and (3) prescriptive data. For descriptive data the teachers answered that they wanted an overview of the course's progress with the possibility to identify struggling students. They wanted grades and evaluations to be included in this overview [11]. For comparative data the teachers wanted the possibility to compare information between students and courses, such as comparing current students' performance with performance in last year's class or a student's performance in different courses and use this information to identify students with difficulties [11]. Lastly, for prescriptive data the teachers answered they wanted historical data to understand possible scenarios in the students' performance and help them rethink and change the lecture plan [11].

Different ways of visualizing data have different uses. For instance, using linear visualizations to represent data from time series, network visualizations to gain deeper insight about how their (the users) peers communicate and activity streams can influence the learners' commitment and awareness [11]. Teachers can use visualization tools to get an overview of the students' progress and knowledge. By having this information, they can adjust the learning process and learning material or change their teaching methods to suit the students' needs. Zheng et al. [4] adds that it should be easy for the instructor to make sense of a visualization, otherwise it might lead to confusion and frustration. Teachers interviewed by Michos et al. [12] and Isaias and Viana [11] also expressed the importance of data visualizations that are easy to understand.

Color coding can be a very useful attribute for a dashboard. It has been used in previous projects to help teachers get information from briefly glancing at the dashboard [10, 19, 22]. Herodotou et al. [10] conducted six semi-structured interviews with teachers, asking about the teachers' experiences using a teacher view dashboard. One of the two most helpful features from the dashboard was the color-coded system which helped identifying the students at risk. The majority of the interview teachers gave attention to the students who were flagged as red (indicating risk), however, one of the teachers found it motivating to look at the students who were doing well and sending them encouraging emails [10]. The teachers also found it useful to see the average score of the current course compared to the previous year as well as average student activity per week, such as clicks.

Khakaj et al. [6] conducted interviews with teachers that revealed that non-visualized student data was more than what they could process or even have need for. This was off-putting for the teachers and led to infrequent use of the software, which emphasizes the need to visualize data rather than displaying it as-is. Bao et al. [20] suggest that by providing multiple alternative visualizations from a more comprehensive analysis of

heterogeneous data, we will be able to better support teachers in applying more effective facilitation approaches. They also found evidence indicating that a visual representation of the social interaction between students can help the teachers provide additional support and encouragement to the students who need it. Dourado et al. [21] found from a design study of a prototype dashboard that both removal of extraneous information and reduction of the dataset were highly favored heuristics.

3 Design and implementation

This section will describe the process of designing and implementing the application. It will go through the requirements, tools used, the design process, descriptions of the existing system (student view and backend), and finally a description of the system architecture.

3.1 Functional requirements

The full list of functional requirements are displayed in Table 2. The requirements have been prioritized with a number from 1-3, with 3 being the highest priority and 1 being the lowest. All functional requirements labeled with the highest priority must be implemented. Those with medium priority should be implemented once the high priority requirements are in place. The low priority requirements can be implemented if time allows it.

The literature found predictive LA, recommender services, and more data to identify students at risk as useful features and data. However, the artifact of this study is limited by the existing self-assessment application for students, from which we have insufficient data to support the implementation of these features. Therefore, these have been disregarded for the development of this artifact.

ID	Description	Priority
FR 1	Should be able to handle announcements	3
FR 1.1	Should be able to post new announcements	3
FR 1.2	Should be able to delete announcements	3
FR 1.3	Should be able to edit announcements	3
FR 1.4	Should be able to limit an announcement to an activity	3
FR 2	Should be able to edit an activity	3
FR 2.1	Should be able to add questions	3
FR 2.2	Should be able to edit questions	3
FR 2.3	Should be able to delete questions	3
FR 2.4	Should be able to create a new activity	3
FR 2.5	Should be able to create new question categories for an activity	1
FR 2.6	Should be able to upload new questions from file (.csv, .txt)	1
FR 2.7	Should be able to see archived activities	2
FR 2.8	Should be able to archive an active activity	2
FR 2.9	Should be able to duplicate archived activity	1
FR 3	Should be able to view visualizations of student data in an activity	3
FR 3.1	Should be able to see individual student progress and performance	3
FR 3.2	Should be able to see an overview of class performance	3
FR 3.3	Should be able to see students progression over attempts	3
FR 3.4	Should be able to see students progression over time	3
FR 3.5	Should be able to see question statistics (activity insights)	3
FR 3.6	Should be able to see a mastery grid for each student in an activity	3
FR 3.7	Should be able to see a list of students with activity performance if there are no categories	2
FR 4	Should display notifications for events that teachers should look at	2
FR 4.1	Should display notification on an activity with issues	2
FR 4.2	Should display notification on students at risk	2
FR 4.3	Should display notification if a question has a high failure rate	2
FR 5	Should be able to authenticate/authorize users	1
FR 5.1	Should be able to log in using username/password	1
FR 5.2	Should be able to create a new user	1
FR 5.3	Should require authorization to post/edit/delete resources	1
FR 6	Should be able to customize the activity view	1
FR 6.1	Should be able to expand a selected visualization	1
FR 6.2	Should be able to hide a selected visualization	1
FR 7	Should see feedback messages (snackbar)	1
FR 7.1	Should see a success message on successful POST/PUT/PATCH request	1
FR 7.2	Should see an error message on error in POST/PUT/PATCH request	1
FR 7.3	Should see form validation	1
FR 8	Should get a confirm action dialog when deleting a resource	2

Table 2: Functional requirements

3.2 Tools

Several tools were used to facilitate the design and implementation of the new dashboard, and they are described briefly in this section. The tools are divided into design tools and development tools. Design tools refers to tools used to create the first three design iterations of the dashboard. The development tools refers to the tools used to create and develop the final prototype.

3.2.1 Design tools

A pen and paper approach using drawing software on a tablet was used to create the first two iterations. This is a quick and easy way to lay out which elements one desires in the design. Since this approach require little effort and time it allows for rapid prototyping. Using a tablet with drawing software also makes it easier to re-draw and re-use elements, further speeding up the process.

Figma was used to create higher fidelity prototypes with interactive elements [24]. While requiring more effort, it can provide highly detailed and interactive prototypes, and styles can be exported to CSS format to be used directly in the code.

3.2.2 Development tools

Development of the code was done in integrated development environments (IDEs). Different ones were used in the development of the SmartU code, mainly IntelliJ [25] and Visual Studio Code [26]. Both offer features and extensions which can make it easier to write clean code.

The institute's GitLab was used to host the code repository. GitLab supports version control with git, which makes it easy to collaborate on programming projects [27].

Docker was used to create virtualized containers [28]. As long as a machine can run Docker it can also run the containers. This makes it easier to develop applications that can run without issue on different hardware.

The application frontend was built using the React library [29] with TypeScript [30]. The backend is a REST API written with NodeJS [31] and Express [32], while the database is written with MySQL [33].

3.3 Design process

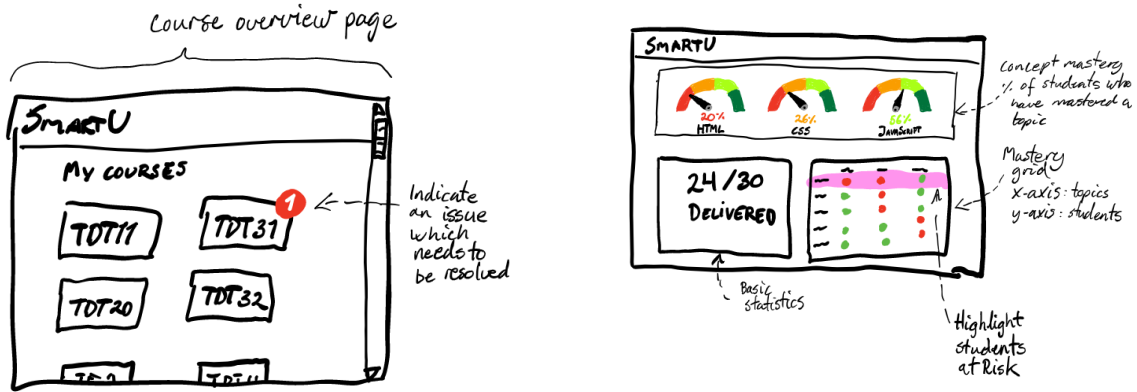
In interface design, an iterative design process starts with the initial planning and requirements, followed by implementation and design. After this, testing and evaluation follows. This leads to new requirements and the process is repeated until one is satisfied and can finally deploy. According to Nielsen [34], upwards of 165% improvement in overall median

usability can be attained from first to last iteration. From one iteration to another 38% improvement can be reached [34]. For the SmartU teacher’s dashboard, the design process was repeated four times, and between each iteration 2-3 people were asked a set of questions relating to the design, which can be found in Appendix C.1.

In this section we will describe the different iterations of the design, starting with the crude paper mockups, to the more detailed higher fidelity prototypes, and finally showing the completed design.

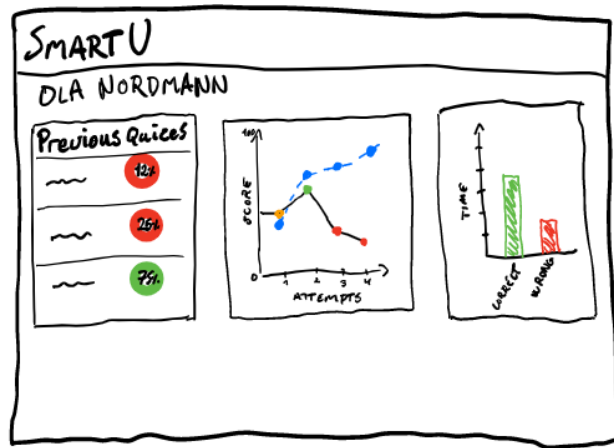
3.3.1 First iteration

The first iteration of the design was a virtual paper mockup which included some of the main elements that were important to the dashboard. Figure 1 shows the three screens which were made for this initial prototype. In Figure 1a we see the course/activity selection screen which was intended as the first screen the teacher would be greeted with. This shows a list of all courses/activities which are available for the teacher, and with a red notification on activities where there are issues that requires the teacher’s attention. Figure 1b shows the dashboard of an activity. This view has gauges for different topics within that activity which shows in % how many of the students that have mastered each respective topic. It also contains a *mastery grid* which has a list of all students in the activity with an indicator for their mastery within each of the topics. The student at the top is classified as being at risk, and is highlighted. The notification on the course selection page and the highlighting in the mastery grid addresses the findings from previous studies [10, 5, 11, 12, 14]. This view also shows some other statistics, such as how many students have submitted a quiz. Finally, in Figure 1c the student history within the activity is displayed. This includes a list of previous quiz attempts and a graph which shows their score over time with a blue line which shows the average for all students.



(a) Activity selection

(b) Activity dashboard



(c) Student activity history

Figure 1: Iteration 1 virtual pen & paper design

Three people were asked questions about the elements of the first iteration. The red notification on the course selection screen was easy to understand for all three users. The gauges for each topic at the top were a bit more confusing, and it was not obvious what they were displaying. On the other hand, the colors appeared to be understandable. Meanwhile, the statistics of deliveries were not seen as very useful. The progression graph and time to answer for correct/incorrect responses on the student view were also difficult for users to understand. Some of the discovered issues were addressed in the next iteration.

3.3.2 Second iteration

Reviewing the comments from users on the previous iteration, and consulting with the findings of the literature review once more, a second virtual paper prototype was made.

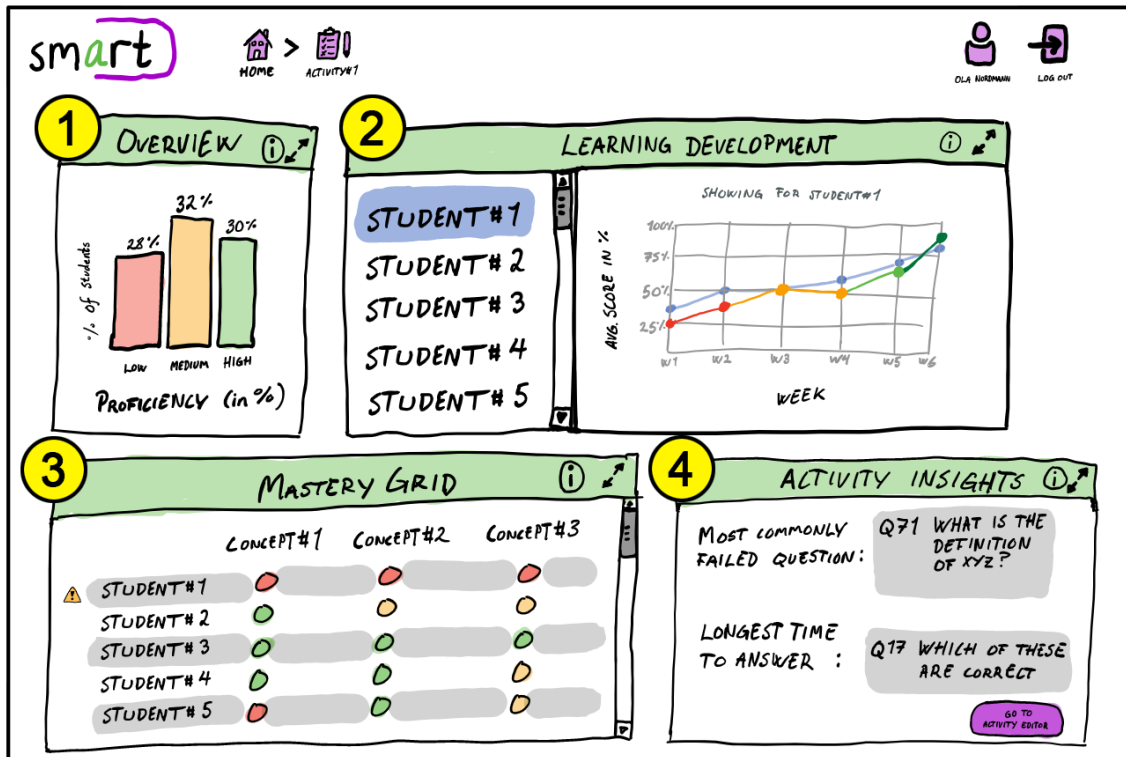


Figure 2: Iteration 2 activity dashboard

In Figure 2 we see the dashboard of an activity. The various data and visualizations are divided by placing them in separate cards/panes. In (1), the overview card shows a simple bar chart of percentage of students within the three proficiency levels, high, medium, and low. The data is collected from the students who have participated in the activity by taking the adaptive self-assessment test, and since this number may change over time it is given as a percentage of the total amount of unique students who have taken the assessment. The learning development graph from the first iteration has been moved from the student details view to the main dashboard view. The learning development card (2) allows the user to scroll through and select a student from the list to view more in-depth information, which here shows their average score per week in percentage. The colored line depicts the score curve of the selected student, while the light blue line shows average score of all students combined. The mastery grid (3) has been kept intact from the previous version. It shows a list of students and displays which concepts (within the selected activity) that they have mastered. This presumes that the activity has questions that are divided into concepts, as this may require an extra large question pool for the activity to work well with the adaptive nature of the assessment. In (4), an overview of some hopefully useful insights about the activity are displayed, addressing some concerns about the less interesting statistics from the previous version. If for example one question is answered incorrectly at a higher rate than others, there may be an error in the assignment of the correct answer.

Each of the cards should be both collapsible and expandable, to take up as little or much

space as needed, as teachers may use the system on different screen sizes. Graphs should be interactive, e.g., so that one can hover the mouse pointer over one of the bars in the bar chart in (1) and be given the exact number of students within that proficiency.

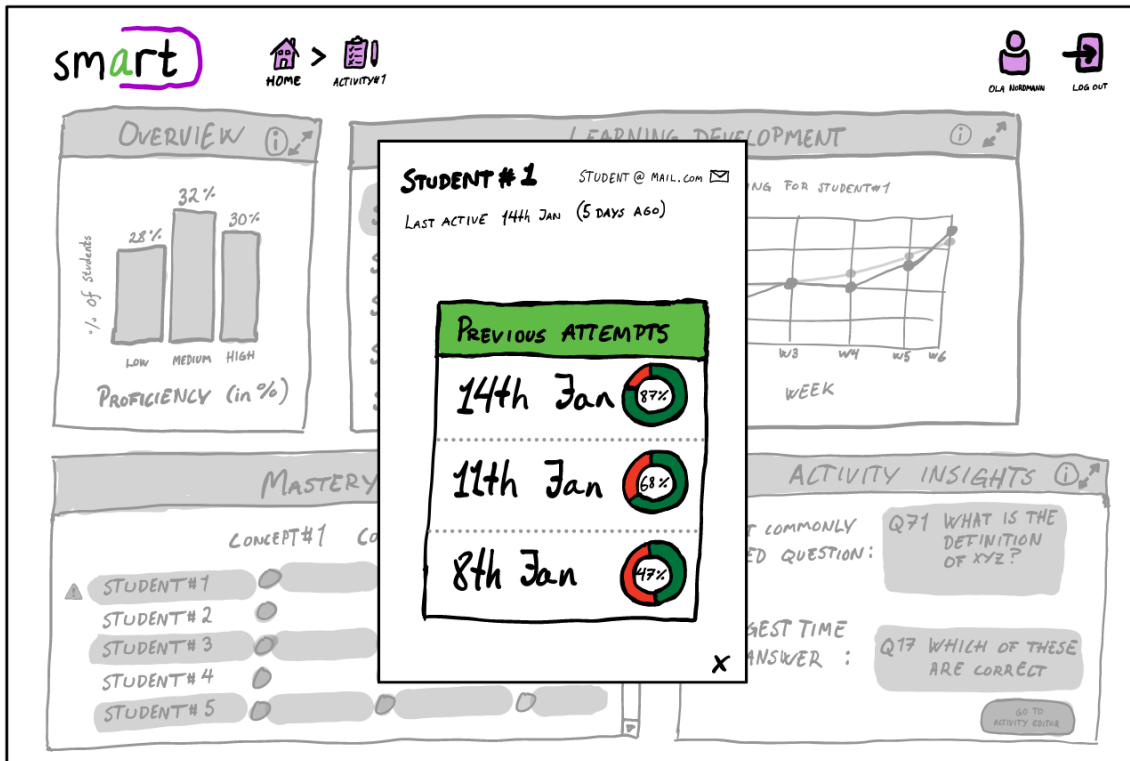


Figure 3: Iteration 2 student details

User tests on this iteration found the overview graph to be understandable. Although the learning development graph was understandable and liked, the users were confused by the two lines and mistook the blue line as representing the selected student because it matched the highlight color around the student in the list. Another issue commented on by one of the users was the issue of finding struggling students when the activity had many students. The mastery grid was easy to understand for both users, and both noticed the highlight on the student at the top. While the activity insights were commented on as being helpful, they were also criticized for being messy with the text outside the boxes.

3.3.3 Figma prototype

After receiving some feedback from users on the previous iteration, the third prototype was created in Figma. As can be seen in figure 4 it contains all the four main elements of the previous iteration, and maintains the additional features in the header such as the breadcrumb and user-related actions.

There are a few changes. The learning development's list of students now has student

names highlighted in the three colors indicating their mastery level, as some users commented that it could be hard to find struggling students in an activity with a large user base. The list is also sorted so that the poorest performing students are placed on top. Legends were also added, as a user found it difficult to understand the difference between the two lines. This should help teachers to focus their attention on those students that may be struggling more than others. The sorting also carries over into the mastery grid, sorted by score in ascending order. The activity insights have also been changed slightly, placing all the information inside the cards, as one user felt it was too detached and this could save space.

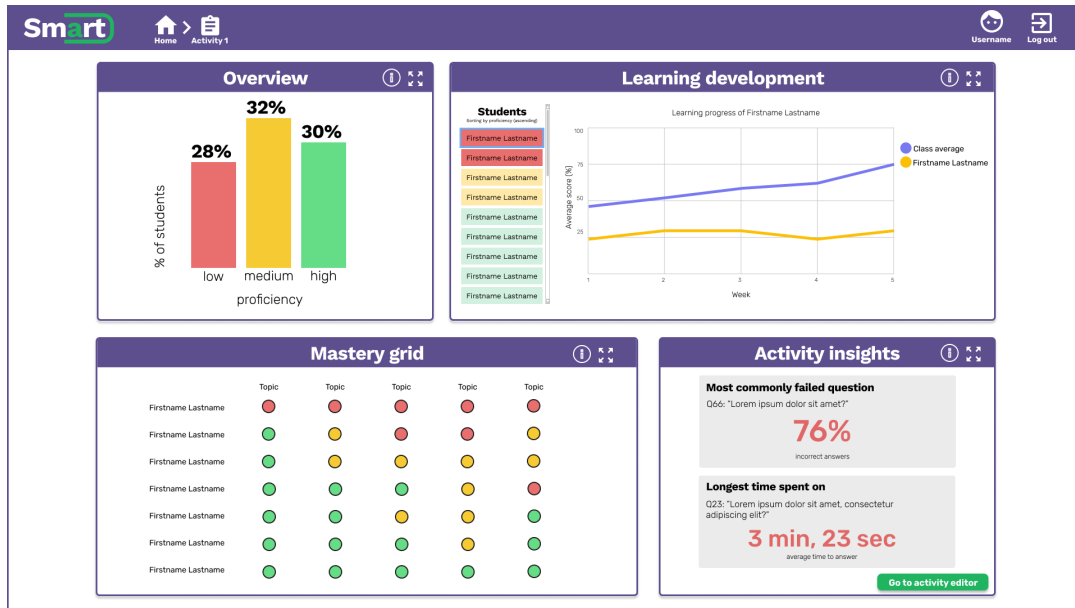


Figure 4: Figma prototype of dashboard

This iteration also included a rework of the front page of the application where the teacher will select the activity, shown in figure 5. A user had commented on the previous iteration of the front page that some teachers may be connected to many activities, which could make it difficult to find the correct activity, especially since there could be many activities with the same name for different semesters. The solution was to only show the currently ongoing activities, and keeping the notification to highlight activities where extra attention may be required. A button to scroll horizontally was added to find old or upcoming activities that may not be that interesting for the teacher to look at. Additionally, information about the number of students in the activity was added to the select activity button.

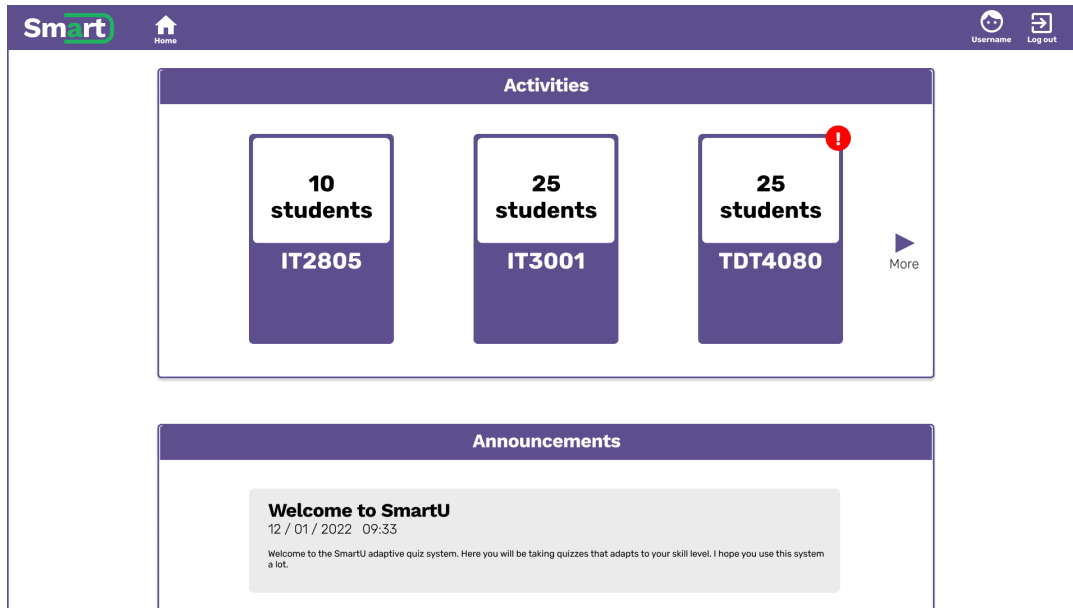


Figure 5: Figma prototype of landing page

The Figma prototype borrows some design elements from the SmartU student view, such as coloring to give a more consistent look across both applications [35]. The main color purple is used consistently throughout the prototype to encapsulate the important elements of the application.

3.3.4 Final design

The final design (see Appendix C.2 for all views) simplifies some aspects of the Figma prototype; the colored trim around the boxes was removed to make the design more modern (figure 6). It should also help users focus better on elements in the view. This point is further accentuated by the extensive use of color coding within the visualizations, where green, yellow, and red are used to indicate the mastery level of students. The final version includes more pages than what was initially prototyped in the previous iterations, as forms for editing and creating questions, activities or announcements are more straight forward than the more novel dashboard view.

To evaluate the Figma prototype, two users were asked to give feedback on the system. One of the users commented that if one wanted more statistics, it would be annoying to have to scroll within the activity insight box. As a result, the activity insights card has been replaced by freely placed cards in the bottom right corner. Removing the outer box gives more space, and gives a more flexible design, as these smaller statistics cards can be moved to other places in the view more freely. The additional space makes it easier to add more of these types of statistics, which has been done in this iteration. It was decided to include some positive statistics as well, such as a counter on login activity and a counter on how many students that has improved since their previous attempt. Since the sorted lists

may give the teacher a negative impression of the activity's status, the positive statistics should work as a counterbalance and make the system slightly more encouraging to use.

In addition to the general color coding and sorting by mastery, notifications have been added to the statistics cards. These are triggered by the values being presented exceeding or falling below a set threshold. They are meant to help focus the user's attention if there are any problematic questions that may require some form of interaction from the teacher. Due to a comment from a user, the data on the longest response time is filtered so that all answers over 20 minutes are filtered out, as it can be assumed that a student has left their computer if they spend more time on a question.

Another feature added due to feedback on the previous iteration was filters. Both the learning progress graph (formerly named learning development) and the mastery grid were given filter options, such as searching for a specific name or the option to change the sorting order. The user features which were present in the previous iteration's header were removed, as the focus of the study did not require users to log in to use the system. Thus, this was de-prioritized and did not make it into the final version.

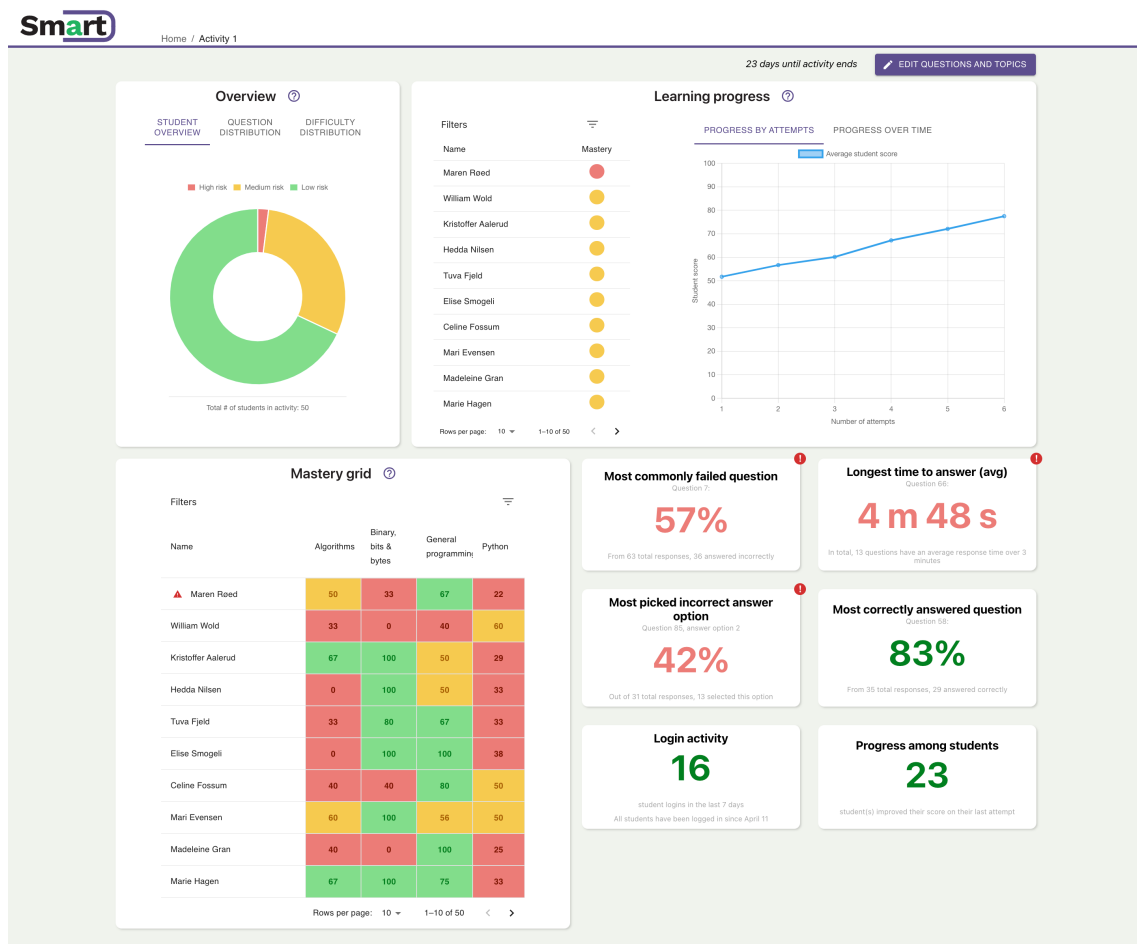


Figure 6: Activity dashboard

Figure 7 shows the landing page of the application. This is where users will begin their interaction with the system. It consists of two main elements, the activity selection card, and the announcement card. The activity selection card has three tabs. The “active” tab is the default, showing activities that are active. The upcoming tab shows activities that have not yet started, e.g., an activity for the upcoming semester. The archive tab shows activities that have been archived or expired. Splitting up the activities is meant to tidy up the view when the user is connected to many activities. The announcement card contains messages that the teacher can create to notify students in the activities. The teacher can write, edit, and delete announcements.

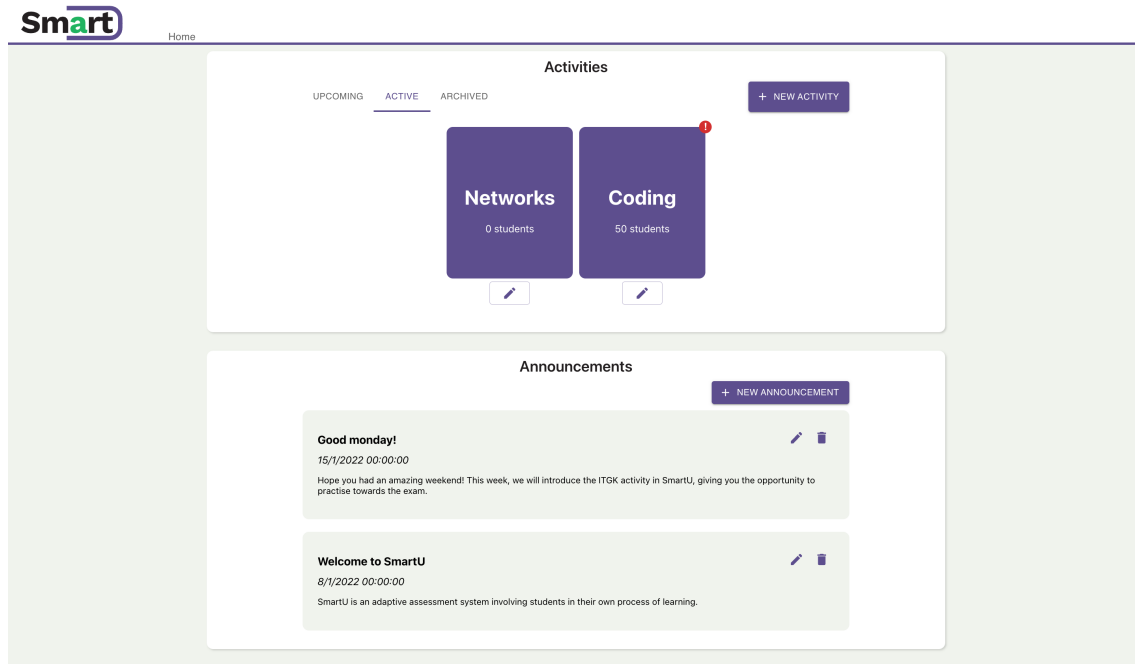


Figure 7: Application landing page

In a system that contains a multitude of questions of varying difficulty it is necessary to have tools to manage them. Figure 8 shows the question management and overview page. This page displays a list of all the questions connected to the activity. Each question has statistics which are continuously updated as students answer questions in the SmartU student application. The statistics include number of correct and wrong answers, the effort (answers where the student spent longer than a threshold value to answer the question), performance (correct answers / total number of answers) as well as the average response time to the question. Questions can be opened and edited from clicking on them in the list. Teachers can create new questions or delete selected questions. They can also upload questions from comma separated value files (.csv) as well as from spreadsheet files (.xls, .xlsx). Questions can also be added with a simple form. There is also the option to edit or add topics.

Double click a question to edit

TOPICS UPLOAD + ADD DELETE

ID	Question	Topic	Difficulty	Correct answers	Wrong answers	Effort	Performance	Average response time
1	Algoritmer er alltid organisert slik at man først	Algorithms	3	48	26	0.946	0.649	2 m 22 s
7	Definisjon av algoritme er i følge teoriboka et i	Algorithms	3	27	36	0.921	0.429	2 m 25 s
8	En algoritme kan beskrives som	Algorithms	1	37	19	0.946	0.661	1 m 56 s
21	Hva beskriver store-theta (Θ) notasjon i forbin	Algorithms	2	30	17	0.894	0.638	2 m 15 s
23	Hva betyr det at en algoritme skal ha en termi	Algorithms	1	27	20	0.894	0.574	2 m 21 s
53	Hva er tidskompleksiteten til denne pseudoko	Algorithms	2	25	19	0.932	0.568	2 m 31 s
54	Hva er tidskompleksiteten til denne pseudoko	Algorithms	2	35	14	0.959	0.714	1 m 52 s
55	Hva gjør denne algoritmen?	Algorithms	1	26	11	0.946	0.703	1 m 8 s
60	Hva kan du si om denne algoritmen?	Algorithms	1	27	6	1	0.818	2 m 25 s
70	Hva stemmer for den naive bondesjakkalgori	Algorithms	1	22	11	0.97	0.667	1 m 41 s
71	Hva stemmer for den smarte bondesjakkalgor	Algorithms	2	19	15	0.941	0.559	2 m 27 s
72	Hva tidskompleksiteten til en algoritme der m	Algorithms	3	15	11	0.962	0.577	1 m 60 s
73	Hva tidskompleksiteten til program med enkel	Algorithms	2	18	18	0.889	0.5	2 m 3 s
74	Hva tidskompleksiteten til rent sekvensielt pro	Algorithms	3	19	14	0.939	0.576	1 m 46 s
77	Hvilke egenskaper har binærseek algoritmen?	Algorithms	1	28	14	0.952	0.667	2 m 10 s
78	Hvilke egenskaper har inntektsorteringsalgor	Algorithms	1	19	6	0.92	0.76	1 m 19 s
79	Hvilke egenskaper har sekvensielt søk?	Algorithms	2	18	13	0.903	0.581	2 m 21 s
82	Hvilken algoritme er dette?	Algorithms	1	20	17	0.892	0.541	1 m 20 s
83	Hvilken algoritme med følgende karakteristikk	Algorithms	3	17	12	0.931	0.586	1 m 30 s
84	Hvilken algoritme med følgende karakteristikk	Algorithms	2	21	9	0.9	0.7	1 m 55 s
85	Hvilken av disse algoritmene er mest effektiv i	Algorithms	2	18	13	0.935	0.581	1 m 59 s
107	Hvordan beskrives ofte algoritmer?	Algorithms	2	16	13	0.931	0.552	2 m 5 s
115	Må stegene i en algoritme være ordnet (følge	Algorithms	1	21	9	1	0.7	1 m 60 s
116	Med at en algoritme skal ha entydige steg me	Algorithms	1	24	12	0.889	0.667	2 m 16 s
2	Bestem True eller False: (A>5 or B==2)	General programming	1	44	19	0.984	0.698	1 m 46 s

Rows per page: 25 1-25 of 122

TO ACTIVITY DASHBOARD

Figure 8: Question management and overview

Creating a new activity or editing an existing one is done with a small form where the teacher must enter a title, a description, as well as a start and end time. Figure 9 shows the form.

Edit activity

Title*

Coding

Minimum 2 characters, 4 characters left

Activity text*

This activity will let you test yourself on questions in Python, Algorithms, Binary and some general programming questions.

Archive activity

Start date

03/29/2022

Start time

11:02

End date

06/03/2022

End time

01:00

CANCEL SAVE AND EXIT NEXT

Figure 9: Edit activity

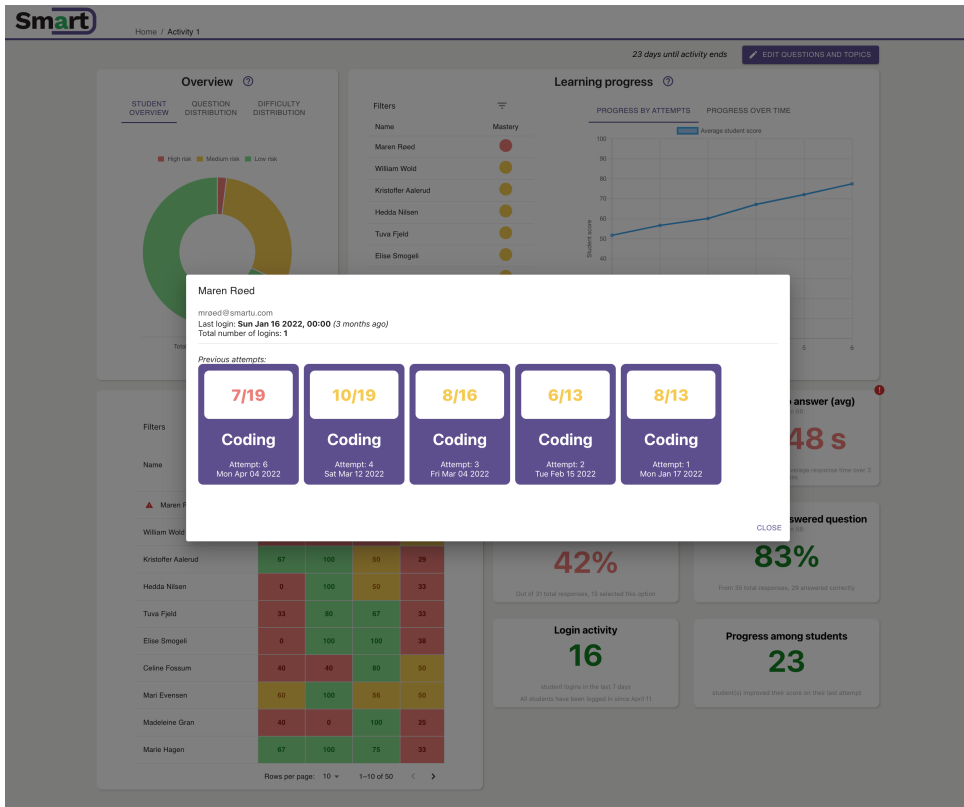
Editing questions does not always require the user to go into the question management page. The statistics showing questions that may be problematic can be clicked on, which opens the question edit form, which is displayed in figure 10. The purpose is to streamline the process of identifying problems and acting on them.

The screenshot displays the Smart LMS interface. In the foreground, an 'Edit Question' dialog box is open, allowing for the modification of a question. The dialog includes a text area for the question, a dropdown for options, and a 'Set correct' radio button. The background shows an 'Overview' section with a donut chart representing student risk levels (High, Medium, Low) and a 'Mastery grid' table showing student performance across different topics.

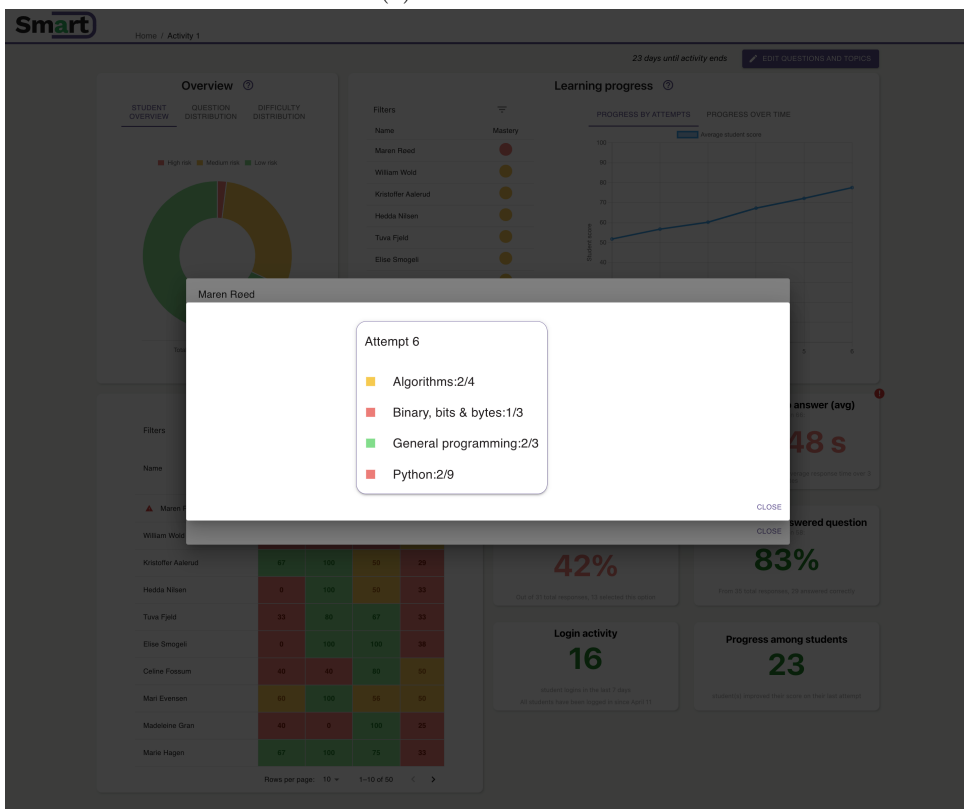
Name	Algorithms	Binary bits & bytes		
Maren Raed	50	23		
William Wold	33	0		
Kristoffer Aalerud	67	100		
Hedda Nilson	0	100		
Tuva Fjeld	33	80		
Elise Smogeli	0	100		
Celine Fossum	40	40		
Mari Evensen	60	100	56	50
Madeleine Gran	40	0	100	25
Marie Hagen	67	100	75	33

Figure 10: Edit question

More details about a student's activity can be found when clicking on them in the mastery grid list. This brings up the dialog shown in figure 11a. Details such as email, date of last login, and total number of logins is displayed. All previous attempts at the adaptive assessment are also shown, again making use of the color coding to indicate performance. Clicking on an attempt brings up information (see figure 11b) about which topics the questions came from with a color code for each topic to indicate how well they did.



(a) Student details



(b) Student attempt details

Figure 11: Student views

Figure 12 displays the final version of the application's navigation and interaction flow. A lot of the interaction in the system keeps the user on the same page, opting for using dialogs opening in the foreground instead of taking the user to a new page.

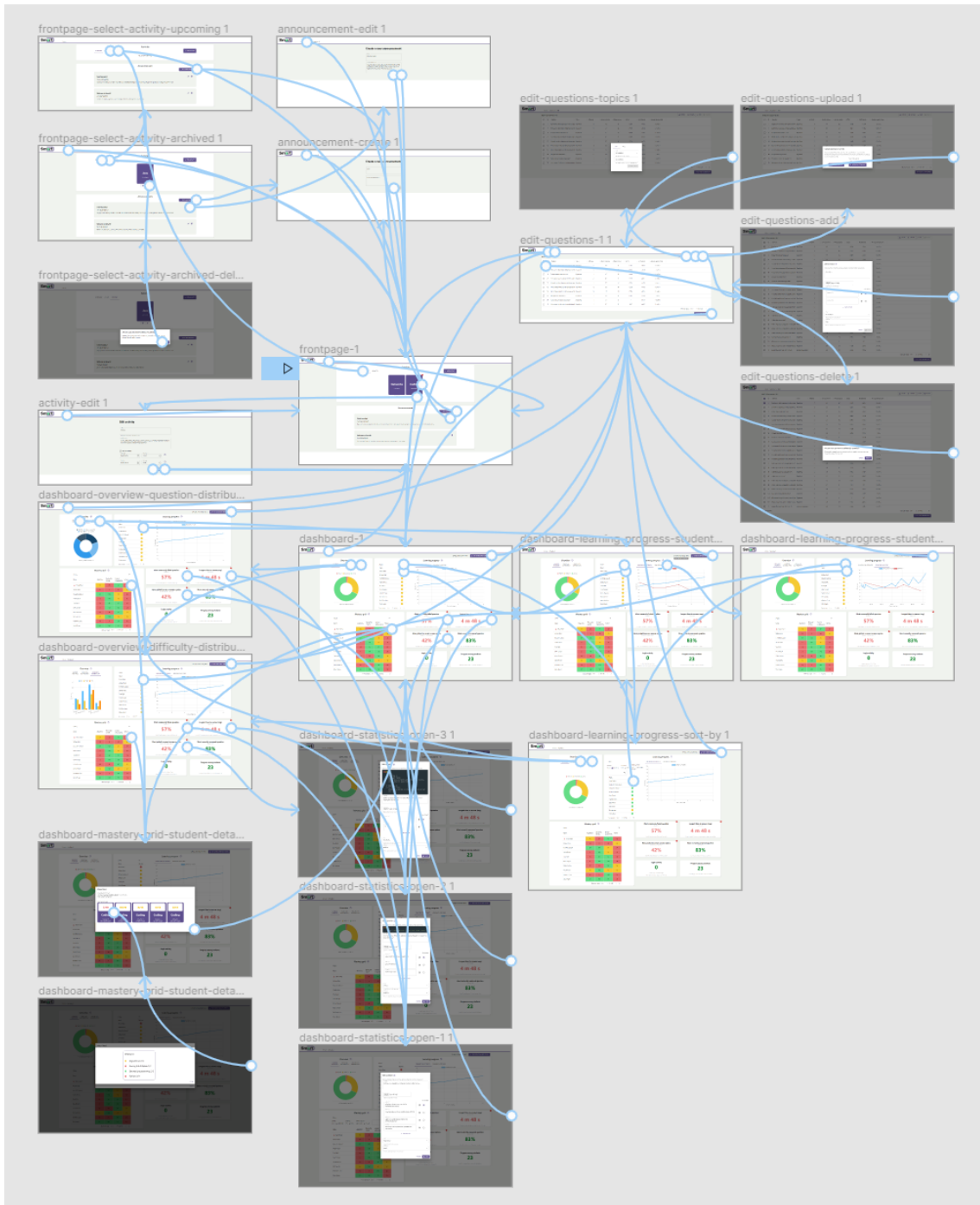


Figure 12: Application flow

3.4 Current system

The pre-existing SmartU system is comprised of a UI for students, which includes dashboard elements and adaptive quizzes. Students' usage of this system generates data that is saved in a MySQL database. Communication goes through a REST API written in PHP. It is the data generated and stored in the relational database that is useful for the teacher UI that is used to produce the information displayed on the teacher-facing dashboard.

3.4.1 System architecture

The system employs a client-server architecture, with the same 3-tiered layers as the student view [35]. The presentation tier is built with React version 17 with TypeScript. The REST API is built with Node and Express on the application tier, while the data tier uses the existing MySQL database. Figure 13 shows the architecture of the new system, also displaying the existing student view created by Westermoen and Lunde [35].

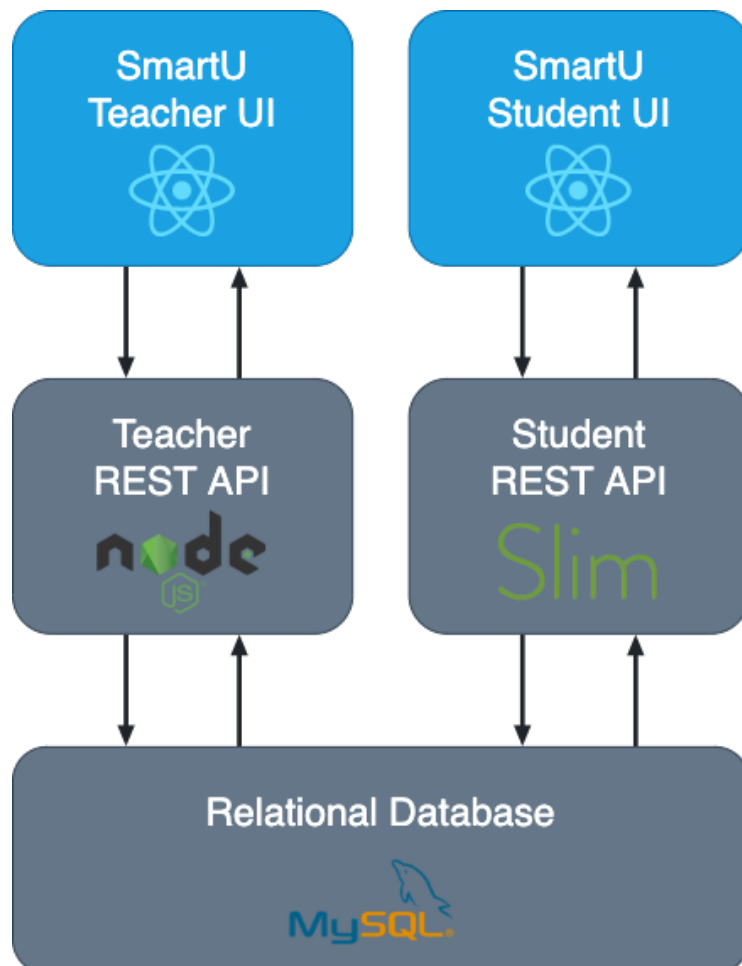


Figure 13: System architecture

Building the presentation and application tiers separate from the student view supports modularity and runs no risk of introducing unforeseen bugs or errors into the existing project. The data tier needs to be shared by both projects, since the data generated by the student view is required to display data to teachers.

4 Methods

This section describes the details of how the study was executed. Subsection 4.1 describes the recruitment process, test setup and the test procedure for the study. Subsection 4.2 describes how and what data was collected from observations, questionnaires and interviews. Finally, subsection 4.3 describes how the collected data was analyzed.

4.1 Study design

This subsection describes the details about how the study was conducted, starting with the recruitment process and the physical test setup, and then how the test procedure including user tests was performed.

4.1.1 Participants

Recruitment of participants for the study was targeted towards teachers, lecturers, PhD candidates, and teaching assistants. Participants were recruited through email, personal messages, an internal bulletin board for study participation at NTNU, as well as word of mouth. The goal was to recruit 20-30 participants to achieve statistically significant results from the quantitative data.

In total, 13 users participated in the study. As seen in Figure 14a there were more men than women who participated, with 10 being male and 3 being female. Ages of participants ranged from 23–57. Teaching experience was divided into teachers/lecturers and teaching assistants, where 5 participants had experience with teaching/lecturing, and 8 only having teaching assistant experience. Three participants had previous experience with Learning Analytics tools or software.

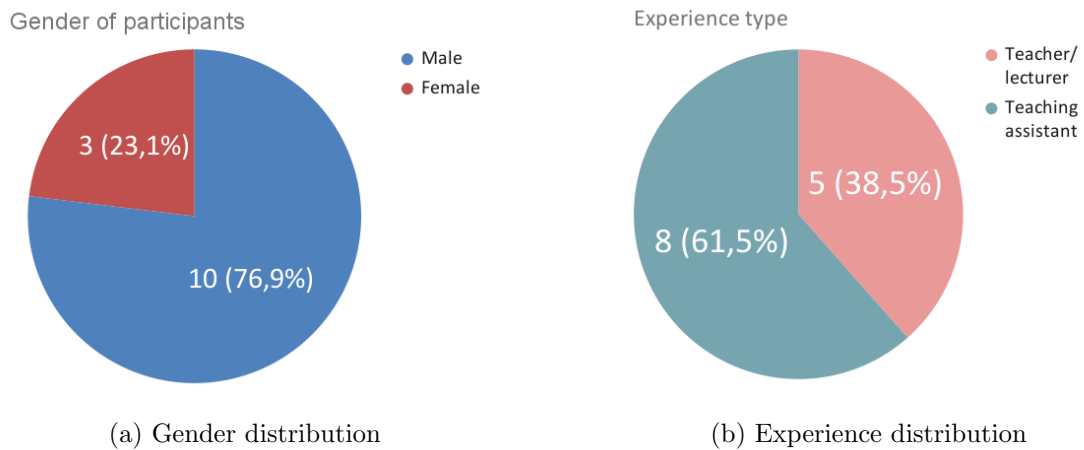


Figure 14: Demographics

N	Min	Max	Median	Mean	SD
13	23	57	28	33.54	11.59

Table 3: Age descriptives

Ten users participated in the interviews. Figure 15 shows that two participants had experience as a teacher/lecturer on university level, and two on high school or primary school level. They will be referred to as teachers in the following sections. Six participants had background as teaching assistants, who will be referred to as TAs.

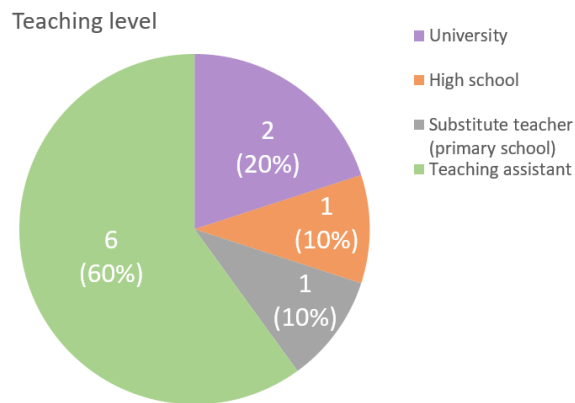


Figure 15: The interviewees' teaching level

4.1.2 Test setup

Two rooms were booked for the study to allow for parallel execution. The setup consisted of a desk, a chair, a laptop connected to an external monitor, as well as external mouse and keyboard. The laptops were loaned from the Department of Computer Science at NTNU,

running Windows 10. The application was running in a Docker container on the laptop, while participants accessed the frontend using the Google Chrome browser. Between each session, the database was reset using an SQL script so that each user interacted with the same initial data.

To create a more inviting environment for the study, snacks and coffee were offered to all participants. Antibacterial wipes were also used between participants and visible so that users would feel more comfortable using the provided equipment.

4.1.3 Test procedure

The test consisted of users interacting with the system after being given a set of objectives, while being under observation of the researchers. For this 15 minutes were allocated. Afterwards, participants were given an online questionnaire to answer which took approximately 15 minutes. Participants could also opt-in for a semi structured interview, which took between 10-20 minutes. In total, sessions lasted between 30 to 45 minutes.

To begin with, the researchers introduced themselves to each participant and offered coffee and snacks to create a friendly environment. Then participants had to sign a form of consent to allow for collection of data from observations, questionnaire and potentially interviews. Then the participants were informed on how the user test would be conducted.

The test began with the reading of a scenario which put the participant in the role of a teacher/TA working in a web development course. After this context had been given, participants were allowed to ask questions if anything was unclear. After the scenario context, participants were given four objectives in order. Participants would complete objectives and inform the researchers when they felt they had completed the objective, so that they could be given the next one. The first objective consisted of finding the correct activity and familiarizing themselves with the dashboard. The second objective had users take action where they deemed it necessary, to test whether or not they would use the dashboard for actions such as reaching out to students or the whole student group, make appropriate changes to questions, or make note of problematic topics. The third objective was to make changes to the activity by extending its duration by two weeks. The fourth and final objective had users create a new activity for an upcoming semester, and upload a file with questions from the test computer.

Upon completion of the scenario, participants were handed an online questionnaire where they were asked questions about the SmartU dashboard. While the participants were answering the questionnaire, the researchers made sure to move so as to not see their answers in real time, thus making the participants more comfortable giving their honest opinion.

Finally, interviews were conducted for those participants that had opted in. The interview was semi structured, and was recorded to later be transcribed. During the interviews the participants still had access to view the application to help them craft more thorough responses.

4.2 Data collection

Data was collected from three separate sources: observations, questionnaire and interviews. For all participants ($N = 13$) the observations and the questionnaire responses were collected, while interview data only came from 10 participants.

4.2.1 Observation

The observations for this study can be classified as participant observation, where the observer takes the role of a complete observer [7]. This means that the researcher making the observations takes no part in the on-goings of the user test, and is silently taking notes on the participants actions. To facilitate consistent notes across participants, the researchers adhered to the observation guide presented in table 4.

Category	Includes	Researchers should note
Test environment	Set up, ambiance, atmosphere	Differences between environment if there are changes between subjects
Emotional state	The emotional state of the test subject (stressed/happy/angry etc.)	Changes in emotional state, any particular aspect of the artifact that probes a specific reaction, is the subject focused and/or motivated
Unexpected events	System bugs/crashes	If anything unexpected occurs note when, how and why
Help	Questions, frequency	What questions does the test-subject ask, and how often do they ask for help
Disturbances	Noises, people entering/leaving the room, phone ringing, etc.	Unwanted noises that can affect the test-subject's concentration and focus on the system

Table 4: Observation guide

4.2.2 Questionnaire

The questionnaire used to capture quantitative data after the user test included four categories of items: demographics, system usability scale (SUS), perceived usability (PU), and perceived ease of use (PEOU). Demographics included gender, age, and type of teaching experience as well as for how long.

The SUS is used to get a subjective assessment on the usability of the system [36]. It contains 10 items and uses a 5-point Likert scale for each, alternating between positive

and negative statements to force participants to consider each statement more carefully [36]. The end result of a SUS is a score ranging from 0–100, where a higher score is better.

Davis defines perceived usefulness as “*the degree to which a person believes that using a particular system would enhance his or her job performance*” [37, p. 320], and perceived ease of use as “*the degree to which a person believes that using a particular system would be free of effort*” [37, p. 320]. Evaluating the perceived usefulness of the system can help answer RQ1 and RQ2.

There were 14 PU items in total, which belonged to five categories of features: overview, learning progress, mastery grid, statistics, and announcements. There were 13 PEOU items, which were also broken up into five categories: content management, highlighting, visualizations, mastery grid, and learning progress. Both PU and PEOU used a 5-point Likert scale ranging from strongly disagree to strongly agree with neutral as the middle point. These can then be coded to numerical form where the lowest is 1 (strongly disagree) and the highest 5 (strongly agree).

The questionnaire can be found in its entirety in Appendix A.4.

4.2.3 Interviews

The interviews were conducted on the same day as the user-test, right after the participants answered the questionnaire, and the interview participation was estimated to take approximately 15 minutes. The interviews were conducted using a semi-structured approach and an interview guide (Appendix A.3). The guide consists of a list of open-ended questions that provides detailed answers for the interviewer to ask. Depending on the answer or contradicting observations during the user-test, some of the questions might require follow up questions. The first question was used as a conversation-starter, and asked the participant about how they felt about using the dashboard. Examples of the predefined questions can be seen in the list below.

- How did you feel about using the SmartU dashboard?
- What was your most preferred visualization in the dashboard and why?
- What was your least preferred visualization in the dashboard and why?
- How do you think you could utilize the information presented in the dashboard for something useful?

4.3 Data analysis

This section describes how the data was analyzed, for both the qualitative data from observations and interviews and for the quantitative questionnaire data.

4.3.1 Qualitative data analysis

The interviews were transcribed before the analysis of the qualitative data started. All interviews can be found in Appendix B. After all the interviews were transcribed, a method inspired by Mayring’s model for inductive category development was applied [8]. According to Mayring’s model of inductive category development, the procedure’s fundamental idea is to develop a defining criterion based on theoretical background and research topic, which decides which aspects of the data are considered [8]. The data is processed through using this criterion, and tentative categories are determined step by step. Those categories are changed, eventually reduced to key categories, and their reliability is assessed as part of a feedback loop.

To interpret the data, categories to describe different recurring themes were at first defined from 3 interviews with interviewees with different backgrounds and viewpoints. Both researchers have a background in computer science and computer engineering, and did partake in this first step where the initial categories were defined, to ensure mutual agreement with the coding rules. After the first step was done together, the researchers split their responsibilities between the analysis of the interviews and quantitative data analysis. The first person to complete their analysis started comparing the data collected from the observations with the findings from the interviews.

As a starting point, 8 categories were created to help group related data from the interviews, as can be seen in Table 5. In addition to a definition per category, each category has "Coding Rules", which are the criteria that must be met for a quote to be placed under a particular category as well as 1-2 examples for each category taken from the interviews. Some of the Coding Rules were changed along the way as potential overlaps in the categories were discovered. When changes had to be made or uncertainty regarding the coding rules arose, the responsible researcher informed and discussed this with their research partner.

Category	Definition	Examples	Coding Rules
Visualizations	Any approach for creating visuals, graphs, or animations to deliver a message is referred to as visualization	<p><i>" I would say that the mastery grid and the filters could be improved a bit. I know that with 50 students it is a bit difficult to explore through all 50 students, but it is useful."</i></p> <p><i>"No, no there was definitely a lot going on but I think it was laid out in a way that I could manage looking at all of it at once, and I don't have to look at all of it at once. But some of those things you do want to look at at the same time."</i></p>	Statements about the visualizations and the dashboard as a whole, except for statements that overlap with any of the other categories.

Highlighted statistics	The six cards at the bottom right with different statistics are referred to as the highlighted statistics	<i>"The only thing that I think may need to be revisited is the fact that zero logins was green, which I think zero logins for a system where the students that if they're supposed to be using the system is bad, and green does not mean bad, green means good so I think that one needs to be flipped."</i>	Statements about the highlighted statistics except statements that overlap with one of the 6 categories in the rows below.
Customization	Customization refers to the user's ability to change values, sorting, sizes, and placement of visualizations	<i>" For sure it can be improved a bit. It could also have some customizable things, but that will be another level. I am happy with the design and it gives the perception that it is very easy to use."</i>	Statements about customization
Privacy	In this case privacy refers to the students' privacy rights, and how much data it is okay to collect and visualize.	<i>"I was trying to figure out how much information I can find about the student. And is it really okay to get students flagged with such a red color?"[translated from Norwegian]</i>	Statements about data collection and the student's privacy
Content management	Content management refers to the user's ability to change, add and remove activities and quiz content	<i>"There were some buttons and menu choices that were a bit unintuitive, it was not clear that you have to double-click it and was a bit unclear that here you have to press if you want to add something new and such. But it's really only about details in the interface, such that could certainly be done with in a relatively short time."[translated from Norwegian]</i>	Statements about editing, adding or removing activities, questions, and categories
Intervention	The information given by the LAD can result in the user wanting to take action/intervene by contacting a student, repeat a part of the curriculum in a lecture or change the curriculum among other things	<i>" I would say this is very powerful and something that is missed in schools and universities, to help teachers understand how or why a task is difficult or not. And to then be able to modify it.", "I want to be able to contact them [...] while looking at their data, and not have to open up a separate thing."</i>	Statements about contacting student, writing announcements, changing the lecture plan or curriculum
Notification	Red dot with exclamation mark on the activity selection page, red triangle with exclamation mark on the mastery grid and red dot with exclamation mark on the highlighted statistics	<i>"I think it's fine. They were a little small, but that might just be because I have old person eyes. But I thought they were fine, and the red one on the purple showed up a lot so yeah it was fine. It told me like «there's a problem here.»"</i>	Statements about notifications

Attitude	Attitude towards the LAD	<i>"I love it. I think you did very good work in terms of the visualizations. I can easily reach what I want."</i> , <i>"If I was actually gonna use it for a class I would want more time, but I don't feel like I would need like a user manual or any hard instructions at all, so it seems pretty accessible."</i>	Statements about how the user feels about using the dashboard and their thoughts about how they are going to use it.
----------	--------------------------	---	--

Table 5: The initial 8 categories

After sorting the interviews according to the rules in Table 5, a majority of the quotes were in the visualizations and intervention categories. To make it easier to analyze the data and see the different perceptions for each visualization, it was decided to divide the visualizations into subcategories for each visualization on the dashboard, including highlighted statistics. The same applies to the intervention category, where the category was divided into two subcategories: contacting individual students and lecture changes. Table 6 shows an overview of the final 7 categories and subcategories.

Main category	Subcategories
Visualizations	- General
	- Highlighted statistics
	- Learning progress graph
	- Overview
	- Mastery Grid
Intervention	- Contacting students
	- Change lecture
Notifications	–
Content management	–
Privacy	–
Customization	–
Attitude	–

Table 6: The final 7 categories with subcategories

In addition to analyzing the interviews, the observations from the user tests were used to supplement the results from the interviews. The observations were sorted by the same categories from Table 6. For this part of the analysis, the main goal was to identify differences between the interviews and observations, such as new findings or contradicting statements.

4.3.2 Quantitative data analysis

The questionnaire responses were collected and transformed. This included changing the 5-point Likert scale responses to numerical values from 1–5. First, information was gathered from the demographic item responses, such as frequency of genders, frequency of teachers and TAs, and experience with LA to give an overview of who the participants were.

The SUS was calculated using Brooke’s method [36]. The odd numbered items (1, 3, 5, 7, 9) was summed up and subtracted 1 point for each item. For the even numbered items (2, 4, 6, 8, 10) the value was summarized then subtracted from 25. Then the result of each were added and then multiplied by 2.5 to give the final score.

The frequency, minimum, maximum, mean and standard deviation were calculated for each of the PU and PEOU constructs. To decide if parametric or non-parametric test would be used, the data was tested for assumptions. Spearman correlation was conducted with all the PU and PEOU categories. Finally, a Mann-Whitney U was carried out.

Table 7 shows the PEOU and PU constructs with definitions and respective item categories. The adopted source for the PU and the PEOU constructs comes from Davis’ definition [37]. Each construct consists of one or more items. Items for each construct were then averaged into one new column per construct in the database.

Construct	Operational definition	Items
Perceived ease of use of Content management (PEOU-CM)	Measures how effortless teachers/TAs believe using the dashboard’s content management features will be	It was easy to change a question in the editor (PEOU-CM1)
		It was easy to add new questions to an activity (PEOU-CM2)
		It was easy to add new announcements (PEOU-CM3)
		It was easy to create a new activity (PEOU-CM4)
Perceived ease of use of High-lighting (PEOU-H)	Measures how effortless teachers/TAs believe using the dashboard’s highlighting features will be	It was easy to understand which questions needed changing (PEOU-H1)
		I could easily understand in which activity there was need for extra attention (PEOU-H2)
Perceived ease of use of Visu-alizations (PEOU-V)	Measures how effortless teachers/TAs believe using the dashboard’s visualizations will be	It was easy to make sense of the color-coding in graphs and visualizations (PEOU-V1)
		It was easy to understand that the visualizations showed the learning progress of the students (PEOU-V2)

Perceived ease of use of Mastery grid (PEOU-MG)	Measures how effortless teachers/TAs believe using the dashboard's mastery grid will be	<p>I could make sense of the mastery grid to see which topics (PEOU-MG1)</p> <p>I could make sense of the mastery grid to see which topics are difficult (PEOU-MG2)</p> <p>It was easy to understand the sorting of the student list for the mastery grid (PEOU-MG3)</p>
Perceived ease of use of Learning progress (PEOU-LP)	Measures how effortless teachers/TAs believe using the dashboard's learning progress graph will be	<p>I could make sense of the learning progress graph (PEOU-LP1)</p> <p>It was easy to understand the sorting of the student list for the learning progress visualization (PEOU-LP2)</p>
Perceived usefulness of Overview (PU-O)	Measuring how useful teachers/TAs perceive the overview graphs	<p>I found it useful to see an overview of the students' potential risk of failing in a doughnut chart (PU-O1)</p> <p>I found it useful to see how the questions is distributed over the different topics (PU-O2)</p> <p>Knowing the topic distribution can be a useful tool to have when adding new questions (PU-O3)</p> <p>I found it useful to see how the questions is distributed over the different difficulties (PU-O4)</p> <p>Knowing the difficulty distribution can be a useful tool to have when adding new questions (PU-O5)</p>
Perceived usefulness of Statistics (PU-S)	Measuring how useful teachers/teaching assistants perceive the highlighted statistics	<p>I found the question statistics to be a useful tool to identify questions with potential issues (PU-S1)</p> <p>I found the question statistics to be a useful tool to identify questions students find difficult (PU-S2)</p> <p>I found the question statistics to be a useful tool to identify questions students find easy (PU-S3)</p> <p>I found the login statistics to be useful to see recent student activity (PU-S4)</p>

Perceived usefulness of Announcements (PU-A)	Measuring how useful teachers/teaching assistants perceive the announcements	I found the announcements to be useful to reach out to students with messages (PU-A1)
Perceived usefulness of Mastery grid (PU-MG)	Measuring how useful teachers/teaching assistants perceive the mastery grid	I believe the Mastery Grid could help me identify students at risk of failing (PU-MG1) I found it useful to be able to click on a student's name to get more details about the student (PU-MG2)
Perceived usefulness of Learning progress (PU-LP)	Measuring how useful teachers/teaching assistants perceive the learning progress graph	I believe I could use the Learning Progress graph to check if the students are on track (PU-LP1) I found it useful to be able to switch between progress by attempts and over time (PU-LP2)

Table 7: Constructs

5 Results

This section presents the results from the qualitative and quantitative data analysis.

5.1 Results for qualitative data analysis

Out of the 10 conducted interviews, 8 are in Norwegian and all Norwegian quotes will be presented as an English translation made by the authors. The original Norwegian interviews can be found in Appendix B. The findings presented in this section are grouped by the categories defined in Section 4.3.1.

To answer the RQs, the results are divided into two groups as mentioned in Section 4.1.1: 4 teachers and 6 TAs. This is done to make sure to address the RQs from a teacher’s point of view and only using the statements from the TAs for comparison or as supplementary data. The Visualizations, Intervention, Notifications and Content management categories provide answers to RQ1 and RQ2, and the Attitude category provides answers to RQ3. The Privacy and Customization categories provide suggestions for future work. Most of the statements from the seven different categories are categorized by positive and negative statements, in addition to other comments, such as suggestions to additional features. The results of the observations have been added where they reveal new information, or where they contradict the results of the interviews.

5.1.1 Visualizations

The Visualizations category is divided into 5 subcategories: General, Highlighted statistics, Learning progress graph, Overview and Mastery Grid. The results will be presented for each subcategory separately.

General	Teachers	TAs
Positive	5	9
Negative	1	0
Definitions	2	0
More aggregated data	1	0
Student background info	1	0

Table 8: General thoughts about the dashboard design

As can be seen in Table 8, many positive statements about the dashboard were given. All participants liked the design and the amount of information displayed. Other positive statements appreciated that the red students are showing at the top as well as being able to sort the lists using the filters.

On the other hand, a teacher found it difficult to understand and find the definition of high risk and low risk. They suggested that information about the meaning of high and low risk should be more easily accessible. They also found it difficult to understand the difference between last attempt and average score. Another teacher, working at a university, wanted more aggregated data and did not want to spend time on individual students. They stated the following:

”And then I think maybe a little more aggregated data, and not for each student. You kind of open the Pandora’s box. You do not have time to sit with individual students [...] it would be unfair to spend more time on individual students. I would rather have more aggregated data at class level.” [translated from Norwegian]

They also suggested that collecting information about the students’ backgrounds, such as previous education or occupation, to be able to search for and identify patterns within students with similar backgrounds:

”It might have been connected to the background of the students in subjects with many students. So that you can e.g. see that those who come from Indøk [industrial management] are struggling with python, and it may be useful information.” [translated from Norwegian]

The observations revealed that most of the participants never tried to hold the mouse pointer over the question marks that contained a description of the visualizations (tooltip).

Highlighted statistics	Teachers	TAs
Positive	4	4
Negative	1	1
Neutral	1	0
Learning objectives	1	0
Question written inside cards	1	0
More questions	1	0
Feedback from students	0	1

Table 9: Statements about the highlighted statistics

The results for the Highlighted statistics from the interviews also gave positive feedback, as can be seen in Table 9, such as easy to understand and being able to edit the questions directly from the main dashboard:

"[...] but just identifying bottlenecks like that is something I've been doing manually in spreadsheets before, so I liked it." [translated from Norwegian]

"The statistics are at least very convenient and the fact that I can even edit directly in the dashboard...! It's pretty smooth. That I don't have to go to the question editor." [translated from Norwegian]

Two negative statements were also given. A TA found it confusing that the card with login information had the color green with zero login, and felt it should have been in the color red when zero students has logged in. A teacher thought that the statistics had too much focus on individual questions. They suggested shifting the focus from individual questions to learning objectives. Another suggestion from the same teacher was to have the question written inside the card, and not only "question 7". A participant wanted to see the second and third most commonly failed question and so on:

"But I would like to click on to the statistics and get more of the worst questions. I can find it in the question editor, but it would have been nice to be able to get it directly from the dashboard." [translated from Norwegian]

A TA wanted more information to identify why a question was most commonly failed, and suggested adding a data point for students to write feedback on individual questions:

"[...] if the student had the opportunity to comment, it could have been a fun data point to have as well, «I think this question was very confusing». Some additional information." [translated from Norwegian]

Observations also revealed that a TA was confused by the difference between most commonly failed question and the most picked incorrect answer option statistics. They felt the wording of these made them difficult to differentiate between. Another TA struggled to understand what the most picked incorrect answer option statistic meant. One teacher was also confused by what the card with progress among students statistic was showing.

Learning progress graph	Teachers	TAs
Positive	4	2
Negative	0	0

Table 10: Statements about the learning progress graph

All statements about the learning progress graph were positive (see Table 10). The statements say that it is very useful to be able to compare individual students with the class average and see if the graph goes up or down at the end.

"[...] it can be a good indication of whether the teaching you do and the plan you have is okay or not." [translated from Norwegian]

"It gives you an idea of what level your students are, and whether there may be some students you should keep an eye on, or whether there are some who find it too easy." [translated from Norwegian]

During the user test, some participants were initially confused by the difference between the two tabs which showed progress by attempts and progress over time. To one TA it was not immediately clear what the points in the graph represented, but this became clear after hovering over them with the mouse pointer. Many of the participants did not try to click on a individual student and only looked at the average score. A TA found it annoying that they could not toggle off the average score line.

Overview	Teachers	TAs
Positive	3	1
Negative	0	0
Bar charts less intuitive	1	0
Eye catching colors	0	1

Table 11: Statements about the Overview visualizations

The results for the Overview can be seen in Table 11. The *Mastery Overview* was the most mentioned Overview visualization among the statements, and all statements regarding the *Mastery Overview* were positive. A teacher commented on the *Question Distribution* and *Difficulty Distribution* to be a little bit less intuitive, but still understandable:

"[...] that is, the bar chart, maybe was a little less... it was quite possible to understand it, but it was not so intuitive to see right away what was displayed." [translated from Norwegian]

A neutral statement from a TA stated that the colors in the visualization was eye-catching and caught the interviewees's attention:

"My gaze was caught by the bright colors at the top." [translated from Norwegian]

During observations, it was noted that many participants regardless of teaching experience did not notice or care to look at the other tabs in the overview card.

Mastery Grid	Teachers	TAs
Positive	1	1
Negative	1	0
Average class topic score	1	2
Compare students	0	1
Show individual questions	0	1

Table 12: Statements about the Mastery Grid

The results for the Mastery Grid can be seen in Table 12. The Mastery Grid received positive feedback on the color coding from two participants. On the other hand, a teacher felt the mastery grid was not consistent with the learning progress graph, and stated that it could be improved:

”And I thought it was also a bit inconsistent in terms of progression. What was in the graph and on these attempts. I think it could have been improved.”
[translated from Norwegian]

Three participants suggested a feature displaying average topic score or a total count of red students in each topic. A TA also suggested adding a function for comparing two or more students would be useful. Another TA wanted to see what questions each student failed on for each attempt:

”Yes, because I see here, there is something with algorithms that has gone wrong, but exactly what questions the student got compared to someone else is something I would like to know. If it was an ID or something like that.”
[translated from Norwegian]

The observations showed that some participants would use the mastery grid to identify difficult topics. One of the teachers found the order of the students in the mastery grid to be confusing, and asked why the student at the top of the list was flagged as at risk when the student below had more red topics. The initial order of students in the mastery grid was not obvious to all; a TA was positively surprised when they found that the results got better the further they clicked through the list. Another TA was confused about what clicking on the different topics on a student row would open, as they expected different details when clicking on the various topics.

5.1.2 Intervention

All participants gave their opinion on contacting students according to the information they found on the LAD. The results in Table 13 represents one answer per participant.

Contact student	Teachers	TAs
Positive	2	5
Negative	1	0
Neutral(depending on the situation)	1	1

Table 13: Statements about contacting students

The results show that the two teachers from primary school and high school level were positive to contacting students. Both also stated that they would not want to contact the students through the system, but rather face-to-face. The two professors teaching at a university had other opinions:

"No, I think the students would have contacted me." [translated from Norwegian]

"I have a bit of a mixed experience with how the students react to being contacted in that way. It is something that everyone knows, that some students like to have a fixed and strict progress plan, while some students like to be more free and travel to Åre on a Thursday. So, it is from individual to individual." [translated from Norwegian]

The student assistants mostly agreed on contacting students according to the information they get from the LAD, except for one participant that would rather prioritize customizing the curriculum. In addition, three of the TAs stated they would like to have the opportunity to contact students directly through the system or to connect it to Outlook:

"But, maybe just some kind of e-mail as soon as I click on Tuva here. Then I get her email address here, but you know there are such links etc. that you can click and your Outlook will open." [translated from Norwegian]

"I tried to click on a person to see if I could send a direct message on that platform, but that may not be the intention." [translated from Norwegian]

"So, I would like an easy or maybe more obvious way to click on a student and to be able to message them directly from where I am looking at the information. I don't want to have to go into a different spot because I want to be able to write that information while I'm looking at – like write her the email while I'm actually looking at her charts and what not compared to everyone."

During the interviews, none of the interviewees mentioned the announcements, but during the observations six participants utilized this feature to contact students. Four TAs and two teachers mentioned or made use of the announcement feature during the user test. A TA would create an announcement and ask those struggling with difficult topics to contact them. Another TA used the announcement to write about problematic questions. A TA used the announcements to write messages of encouragement for students who had improved. The final TA was initially going to use the announcements to reach out to struggling students, but felt it was rude. The same TA also felt that contacting all struggling students through email would be a lot of work. A teacher also mentioned using the feature if there were any questions that many struggled with or if they identified a mistake in a question. The second teacher said they would use the announcements to reach out to those not using the quiz.

Change lecture	Teachers	TAs
Positive	3	7
Negative	0	0
Ignoring skilled students	1	0
Organize by learning objective	1	0
Connect to other tools	1	0
Individual feedback to students	1	0
Correlate with other data	0	1

Table 14: Statements about changing lecture plan, curriculum etc.

When asked about how they would use the information displayed on the dashboard, different opinions and suggestions were shared, as can be seen in Table 14. Most statements were positive in the context of repeating content, changing the progress plan as well as changing and adding questions to the activity. Two of the three positive statements from the teachers were the following:

"The most obvious thing is... it is to be able to identify i.e. topic details for repetition and for extra exercises and in a slightly larger time perspective progress plan for the semester and even larger time perspective, i.e. the syllabus itself." [translated from Norwegian]

"I can take on a topic, and it must be because either a group of the class or the whole class finds it difficult or that there are individuals who need extra adaptation or extra resources, other ways of understanding the content." [translated from Norwegian]

Three of the six positive statements from the TAs were about making themselves more accessible to the students that might be struggling, without contacting them, to make it easier to ask for help. One of the statements were the following:

"And as a TA, it may be that you may be trying to reach out to those you know - or trying to make yourself extra accessible to the people you see may be struggling a bit in the course."[translated from Norwegian]

The remaining four of the positive statements from the TAs were about changing the difficulty levels on the questions across the different topics, as well as adding new questions/topics, according to the information found on the dashboard. One of the statements can be seen here:

"Yeah, for like orchestrating future learning tasks based on all of the statistics about the questions. And also knowing what the distribution of the topics, the

themes and the difficulty level across those were. By looking at those things you can see this is what I've laid out for the kids, and this is how they're doing and what questions are the hardest and potentially why. So by combining those and also knowing what the class average is looking like, I think that I could use that lots to kind of gear what the next content should be or if I should change the content that I have provided."

Some statements were more concerning, such as the LAD can make the skilled students invisible and only give the struggling students attention:

"I think the benefit is that you catch students who are not doing well. I think a risk could be that it gets very numerically controlled and that students disappear." [translated from Norwegian]

A teacher came with a suggestion about having the activity organized according to learning objectives instead of single questions and attempts for individual students as well as connecting it to other learning tools:

"If I could have similar tools organized by learning goals, that it was not about questions and the number of questions, but that I could see which learning goals are easiest to understand. Because I do not think I would have it per student, but rather in general about the subject. And then I could have connected it to a tool such as Mentimeter." [translated from Norwegian]

Another teacher suggested the opportunity to give feedback to students through the system:

"It might have been nice with a feedback function. It could be an idea here. It can be useful if the students do not see me in a few weeks, feedback that not everyone sees, but feedback to each individual student." [translated from Norwegian]

A TA suggested you should correlate the information from the dashboard with other data, such as lecture attendance, before changing the lecture:

"[...] you have to correlate it with other data also for sure. That is, attendance percentage in the lectures and such, to find why and where the problem lies." [translated from Norwegian]

The observations were consistent with what the participants said during the interviews.

5.1.3 Perceptions towards notifications and highlighting

Notification	Teachers	TAs
Positive	4	6
Negative	0	0
Mixed feelings	1	0
Too small	0	1
Prefers mastery grid	0	1
Option to turn on/off	2	2
Different types of icons	1	1

Table 15: Statements about the notifications

During the interviews, some of the interviewees mentioned the notifications in a positive way before any questions about the notifications were asked. On questions about them, almost all were positive, while some were critical/neutral as well as some suggestions of improvement, as can be seen in Table 15. The answers among the teachers and TAs were very similar, so the groups did not differentiate. Among the positive statements were the following:

"I saw that it was an exclamation mark, and just thought that «ok there's something that needs my attention»." [translated from Norwegian]

"Yes, I think it would be useful. Especially if it's to be used over a longer period, and you only log in occasionally." [translated from Norwegian]

"Yes, I think it's good. It's important that it is not too much, but if there are 3 such red dots, it's fine." [translated from Norwegian]

Other statements were more unsure or neutral. One participant had mixed feelings since they did not look at the elements with notification until last and said:

"There was also a notification on the front page, which tells me that there are some updated values. Something like that is very nice to be told about. At the same time, it should be... It can depend on the person, but it should have been the first thing I checked out if I thought it was good." [translated from Norwegian]

A TA commented that the notifications were small:

"I think it's fine. They were a little small, but that might just be because I have old person eyes. But I thought they were fine, and the red one on the purple showed up a lot so yeah it was fine. It told me like «there's a problem here»"

Another TA commented about using the color coding, the red ones, in the mastery grid instead of the student notification:

"It's almost better to just look at the red. She who is next under the student at the top may struggle more since she has 3 reds. I think maybe the color code is the best here, so the triangle may not be necessary." [translated from Norwegian]

A few suggestions of improvement were also given. The possibility to be able to turn off the notification were discussed, and multiple of the participants, both teachers and TAs, liked the idea of being able to turn off the notifications. The participants agreed on that they would like to have the notifications turned on for the first weeks or months of using the system, but some would like to have the possibility to turn them off later. Other suggestions such as having different types of icons as well as other colors were also made by a teacher and a TA, as can be seen in the following quotes:

"Yes, it's better to have them yes. I mean if these systems help me sort out. And especially to understand that it's new since last time. But the question is, is it new or is there something to tackle? There could be different types of notification icons, whether it's just new or something extra." [translated from Norwegian]

"Yes, I think that was very useful, also the red triangle. Maybe you could also have a green exclamation mark on the things that stand out as very easy as well. Because it's easy to see red, red, red, scary! But then you may not see the questions and things that do not bring any progress or development then."[translated from Norwegian]

Observations showed that all participants were able to quickly identify the correct activity, which was marked with a red notification. All participants with the exception of one TA looked at the highlighted statistics. This may not be attributed to the notifications as most of the participants also looked at all the other main elements on the dashboard. A majority of users also made note of the student which was marked with a notification in the mastery grid.

5.1.4 Content management

Content management	Teachers	TAs
Positive	3	0
Negative	2	0
Placement of upload picture	0	1
Feedback when upload is complete	0	2
Personal notes	0	1

Table 16: Statements about content management

The interviews revealed some thoughts about the system’s content management, as can be seen in Table 16. The possibility to upload a file with questions in a basic format such as .csv was well received by a teacher, but they did not like that you have to double click to edit a question:

”There were some buttons and menu choices that were a bit unintuitive, that it was not clear that here you have to double-click.”[translated from Norwegian]

Another teacher enjoyed the ability to sort the questions and statistics in the question editor. On the other hand, another teacher found the question editor to be clunky and difficult to sort. The TAs did not have any positive or negative feedback, but came with a few suggestions and neutral statements. A TA meant the button for upload image should be below the image itself. Two suggestions from the TAs were to give feedback on completion of file upload. Having notes, to write down a to do list, inside the LAD was also suggested.

Observations on usage and comments on the content management features confirms that participants struggled with some clickable areas and menu options. Several participants struggled to open up the edit question form in the list of questions, specifically the double-clicking. Some users did not realize or spent a long time to find that they could open questions from the highlighted statistics. Some also struggled to find where to edit an activity. There were also positive remarks on content management features; a TA exclaimed that editing questions right from the highlighted statistics was very smooth. A teacher thought the edit question page with the list of questions was nice, and liked that he could manage topics. A TA noted how he would use the question list with statistics to improve the learning content, by deleting questions with more incorrect submitted answers than correct.

5.1.5 Privacy

Privacy	Teachers	TAs
Unnecessary information stored	1	0
Aggregated data instead	1	0
Reservation against data collection	0	1

Table 17: Statements about privacy

Among the 10 interviewees, 1 teacher and 1 TA were curious about the privacy policy and how okay it is to collect this much information about individual students (Table 17). The teacher meant unnecessary information is stored about the students, and that aggregated data could be just as useful but less intrusive. The TA asked if there is possible for the students to reserve themselves against sharing information, and that it could be a useful feature to have:

”Is it possible, for example, to have a project where the students have an option to reserve themselves against the teacher being able to see what they have achieved? I would like to see that. If the teacher discovered that it makes the students stressed and if they feel monitored, it would have been nice to have a functionality to turn it off.” [translated from Norwegian]

No comments regarding privacy were noted during the observations.

5.1.6 Customization

Customization	Teachers	TAs
Change threshold values	1	1
Customize size and placement	1	0
Range of time	1	0

Table 18: Statements about customization

Table 18 shows the results for the statements about customization. Out of the 4 statements about customization, 2 wanted to be able to change the threshold values, such as the thresholds for green, yellow and red. A comment about customizing the sizes and placements were also made, as well as a suggestion to give the user a range of time which they can select:

“Yeah, maybe this is even more difficult, but having a range of time which I can select. I don’t know. The first month related to the last would be more

useful for me to better understand the average of the score. It is even more complicated, but I think if you design in a good way it is very useful to have the possibility to select a range of time.”

Sorting	Teachers	TAs
Last attempt	0	3
Few last attempts	1	0
Last attempt or few last	1	0
Few last attempts or average	0	1
Average score	0	1
Unsure	2	0
Range of time	0	1

Table 19: Statements about how the student scores should be sorted

The results shown in Table 19 represents the participants thoughts on how the student scores should be sorted, one answer per participant. Two teachers were unsure about how they would prefer to have the lists sorted. Three TAs thought using the last attempt is the way to go, while the rest of the participants had all different opinions.

5.1.7 Attitude towards using the LAD

The participants attitude refers to their feelings and motivation to use the system, as well as usefulness and how they would want to use it and how often.

Attitude	Teachers	TAs
Positive	4	9
Negative	0	0
Neutral	3	0
Would use	2	6
Would use sometimes	1	0
Would not use	1	0

Table 20: Attitude towards using the LAD

As can be seen in Table 20, most of the interviewees gave positive statements. In total only 4 out of 16 statements were neutral, and all interviewees had at least one positive statement. All participant gave one statement each about if they would use the system or not. While all 10 participants agreed on that the amount of information displayed on the dashboard was a good (not too much or too little), a few said that it will take some time, maybe more than the 15 minute user test, to familiarize themselves with the system to use it more efficiently. One statement among the teachers summarizes this:

"I think it was very helpful. It gave a lot of insight. A lot. So I think... it took me a while to understand what to look at and how to interpret it. But eventually I got... it was very easy to use. How the students are doing I mean." [translated from Norwegian]

Some participants also stated that the system had functionality and information they have never seen in other systems. On the question about if they would use the system if they had the possibility another teacher said:

"Yes, everything must be seen in a context, but I certainly see it as possible because there is some functionality here that I have not seen in other systems that I really wanted." [translated from Norwegian]

Another teacher did not have the same point of view as they did not find it useful for their type of teaching, and stated the following:

"We use very few quizzes, at least not as part of the grade. We use a lot of feedback, i.e. written feedback, so for me it would not be very useful." [translated from Norwegian]

The last teacher did find the LAD useful, but would only use it once in a while to create variety:

"[...] I think it's definitely a useful tool, but I probably would not use it very often, but I would probably have it as something I use once in a while to create variety in my teaching. Plus that if there are things I want to keep track of more than... or easier to keep track of." [translated from Norwegian]

The statements among the TAs were much more similar to each other compared to the teacher responses. They were all positive, and they all said they would use the system. Some of the comments were as follows:

"No it was good, nice, it was easy to use, intuitive. It had, yes, pretty much the information I think is necessary." [translated from Norwegian]

"It was good. It was very easy to get used to. It was not too much. I did not feel it was anything more than it needed to be. It was clear. Very nice. Yes, it was a pleasure." [translated from Norwegian]

"As long as my content was online, then I don't see why I wouldn't use it. I would use it. All the time. Depends on what my curriculum is, but I would use it as often as it was applicable to the content that I was serving to my students. I think it's awesome."

Only one of the interviewees, a TA, hesitated while answering on the question about how often they would use the system. The reason for the hesitation was because they are skeptical about using IT in teaching, as explained in the following statement:

"I'm a little skeptical about using IT in teaching. I don't think one should use IT when there are good alternatives. But this is like there are no good alternatives, which I believe in. So if I was a teacher, I think I would have liked to have used this." [translated from Norwegian]

5.2 Results for quantitative data analysis

5.2.1 Testing for assumptions

The assumption of linearity is met, as well as the assumption of independence of observations. Normality was tested using the Shapiro-Wilks normality test. Out of the 10 constructs of PU and PEOU, only 3 yielded a p-value > 0.05 , and as a result non-parametric tests were used.

5.2.2 Descriptives

Descriptive statistics for the PEOU and PU constructs are presented in Table 21. The results show an overall high score towards both the perceived ease of use and the perceived usability of the application. The mean score for every construct were positive (> 4.00). None of the PEOU constructs have a minimum score below neutral, which indicates that the system was easy to use. The PU constructs were also mostly high, except for statistics (PU-S) which had a minimum of 2.75, which could suggest that this feature was more divisive.

One potential issue of the questionnaire became apparent when participants had not noticed the feature or element questioned about in an item, as there was no option for this besides "Neutral". This could contribute to skewed results for some elements.

Factors	Min	Max	Median	Mean	SD
PEOU-CM	3.50	5.00	4.50	4.40	.582
PEOU-H	4.00	5.00	4.50	4.46	.431
PEOU-V	3.00	5.00	4.50	4.46	.660
PEOU-MG	3.00	5.00	4.00	4.10	.644
PEOU-LP	3.50	5.00	4.50	4.58	.449
PU-O	3.40	5.00	4.60	4.40	.565
PU-S	2.75	5.00	4.50	4.31	.639
PU-A	3.00	5.00	5.00	4.31	.855
PU-MG	4.00	5.00	4.50	4.62	.363
PU-LP	3.00	5.00	4.50	4.38	.712

Table 21: Descriptives for PEOU and PU items

5.2.3 Correlation

Table 22 displays the correlation between the perceived usefulness and perceived ease of use items from the questionnaire. Correlation is deemed significant at the $p < 0.01$ and $p < 0.05$ level.

A strong positive correlation was found between the perceived usefulness of the overview graphs (PU-O) and the perceived usefulness of the learning progress graph (PU-LP) ($\rho = .89$, $p = .000$, $n = 13$). Strong positive correlation was also found between perceived usefulness of the announcement features (PU-A) and the perceived ease of use of the visualizations (PEOU-V) ($\rho = .70$, $p = .008$, $n = 13$). Perceived usefulness of the mastery grid (PU-MG) and the perceived ease of use of the visualizations (PEOU-V) also had a strong positive correlation ($\rho = .79$, $p = .001$, $n = 13$).

Measure	PU-O	PU-S	PU-A	PU-MG	PU-LP	PEOU-CM	PEOU-H	PEOU-V	PEOU-MG	PEOU-LP
PU-O	1									
PU-S	.18	1								
PU-A	.47	.55*	1							
PU-MG	.11	.40	.41	1						
PU-LP	.89**	.06	.50	.03	1					
PEOU-CM	.45	.52	.44	.20	.53	1				
PEOU-H	-.14	.66*	.04	.41	-.13	.47	1			
PEOU-V	.30	.53	.70**	.79**	.16	.33	.29	1		
PEOU-MG	.27	.43	.15	.56*	.27	.55	.66*	.47	1	
PEOU-LP	.57*	.30	.47	.13	.64*	.55	.02	.16	.10	1

Note: *Correlation is significant at the 0.05 level. **Correlation is significant at the 0.01 level.

Table 22: Spearman correlation matrix

5.2.4 Mann-Whitney U test

As the assumptions were not met, the non-parametric Mann-Whitney U test was used to compare the sentiments of the teachers and the TAs. Differentiating between teachers and TAs is useful as they have different methods and objectives. With a larger sample size, one could make a differentiation between teachers at different educational levels as well. Table 23 shows the results of the Mann-Whitney U test. There is a significant difference in median between teachers' and TAs' responses to item PEOU4 about the perceived ease of use of the mastery grid ($W = 5.0$, $p = .031$).

Factor	W	p	Teachers		TAs	
			Mean	SD	Mean	SD
PEOU-CM	14.5	.452	4.30	.542	4.47	.632
PEOU-H	13.0	.313	4.30	.448	4.56	.417
PEOU-V	13.0	.313	4.30	.570	4.56	.729
PEOU-MG	5.0	.031	3.60	.366	4.42	.586
PEOU-LP	23.5	.634	4.70	.274	4.50	.534
PU-O	15.5	.553	4.24	.670	4.50	.513
PU-S	16.0	.603	4.10	.802	4.44	.530
PU-A	14.0	.375	4.00	1.00	4.50	.756
PU-MG	14.0	.381	4.50	.354	4.69	.372
PU-LP	14.0	.395	4.20	.760	4.50	.707

Table 23: Mann-Whitney U test results

5.2.5 Usability

The SUS was first calculated for each response, then the descriptives shown in Table 24 were calculated from these scores. The mean score of 85.38 is considered an excellent score [9]. Even a low sample size of 8-12 users can give confident results [9].

N	Min	Max	Median	Mean	SD
13	70.00	100.00	87.50	85.38	11.27

Table 24: Result of SUS items

6 Discussion

In order to assess the newly developed LAD, three research questions were created. The teacher-facing LAD includes multiple ways to manage quiz content and see statistical data from students taking quizzes in activities, as well as writing announcements to the students. While it is possible to measure a student's performance by looking at correct and wrong answers in a quiz attempt, it is not that easy to measure how well the LAD supports a teacher in reaching their objectives. Because of this, the research questions will be addressed by interpreting the results of the qualitative and quantitative data analysis and discussing possible reasons and solutions in the following subsections.

6.1 Effectively supporting teacher intervention

The first research question we presented in this thesis was: "How can an LAD effectively support teacher intervention?" The findings of the interviews indicates that there is a difference in how teachers use the LAD for intervention based on what level of education they teach at. While university lecturers were skeptical to contacting individual students, those teaching at lower levels of education were more positive, albeit they preferred doing this type of intervention in person. Another form of contacting students indirectly was the announcement feature, which only one teacher made use of as a form of intervention. From the literature it has been shown that primary school teachers make use of a pedagogical intervention at individual and class level to a large degree [14]. TAs were more positive than teachers to contacting individual students. TAs already have tasks that involve a lot of one-on-one interaction with students, which is an objective that this LAD supports. From a teacher's perspective this appears to depend on what level of education they teach at. The Mann Whitney U test found a significant difference between teacher and TA sentiment towards the perceived ease of use of the mastery grid. In connection with the qualitative data this might further implicate the individual versus class-level difference that exists between teachers and TAs. Taking into considerations what some teachers had to say about visualizations, where it was suggested to display aggregated data due to the time constraints and not being able to help individual students, one could argue that the dashboard should offer different visualizations to support intervention based on the teaching level of the user.

There was a strong positive correlation between the perceived ease of use of the visualizations and the perceived usefulness of the announcement feature. The same was true for the correlation between the perceived usefulness of the mastery grid and the perceived ease of use of the visualizations. This could be an indicator that those that found the

visualizations easy to read also found this form of intervention more useful.

6.2 Effectively supporting teachers in adapting their lesson plan and learning format

The second research question we stated in this thesis was: "How can an LAD effectively support teachers in adapting their teaching, lesson plan and learning format?" To address and discuss this research question, we will take a look at the results from the visualizations, content management and customization and compare it with findings from previous studies.

6.2.1 Perceptions towards the different visualizations

While the results show that the participants found the amount of information and visualizations on the LAD appropriate, some participants had difficulty understanding what the color coding and visualizations meant. This was shown mainly in observations, but also in an interview. Color coding has been found as a useful and helpful attribute for LADs in previous studies, helping the user get information from a brief glance at the dashboard [10, 19, 22]. The color coding in SmartU is only described using the words "low risk", "medium risk" and "high risk", which a few participants asked what meant during the user tests, and also mentioned by a teacher during the interviews. While all visualizations had a question mark with a tooltip describing its visualization, few participants hovered the mouse pointer over the tooltip. A solution for these misunderstandings may be to have a short tutorial the first time a user. This can ensure the users understand the LAD correctly, and thus make appropriate evaluations to reach their objectives.

Depending on the class size and how the teachers teach can also impact what the LAD should display. From the interviews we have seen that more aggregated data on class level and less data on individual students can be more appropriate for lecturers working at universities, since they often have bigger classes and less time to spend on individual students. In comparison, the teachers working at primary school and high school will have more time to tailor some of the teaching to individuals or groups. We can also see from the interviews that it could be interesting for teachers at universities to see statistics grouped by different backgrounds, by collecting data about students' backgrounds. This is similar to the findings from previous studies, where a comparison functionality was found to be a useful feature for teachers in a study on an LAD with this functionality implemented, while it was a much desired feature in a study without it [10, 11]. These findings tell us that for the LAD to be useful for different teaching methods, it could be a possibility to choose whether the displayed data is for individual students or if it is more general data on class level. The findings also tell us that adding a comparison feature for teachers to use when they want or need to can be helpful to identify patterns in different groups, such as students with similar backgrounds or students at high or low risk.

The highlighted statistics received positive feedback from the participants. The possibility

to see the questions with most wrong answers or longest time spent and directly edit them from the dashboard was appreciated by most participants. This is in accordance with the group *challenges*, one of the four groups made by Alevan et al. [15], which are different groups of analytics that can be useful for adapting tutoring. The group *challenges* refer to easily identifying the challenges the students face. The findings from our study also found that the teachers stated that the highlighted statistics would save them time compared to using the question editor. Still, there is room for improvement to further assist and help the teachers reach their objectives. A possible change according to suggestions from the study is to give the teacher an option to have the statistics grouped by learning objectives defined by the teacher. Another possible adjustment to help the teachers more effectively evaluate the questions could be to implement the possibility to click next on most incorrect answered question and get to see the second most incorrect answered question and so on.

The learning progress graph was easy to use and understand according to all participants. Some of the statements included that it gives an indication of whether the lectures and the lecture plan is okay or not. This is similar to the findings by Xhakaj et al. [16], where they found that a visualization of student progress over time displayed as a scatter plot was helpful to teachers as they prepare for a lecture. While all statements from the interviews in this study were positive, we noticed during the user tests that many of the participants never tried to click on a student name, resulting in only seeing the average class score. Even though it was observed that the participants hovered over the names, some of them still never clicked on any of the names. Although the average SUS score could be classified as excellent, it does not mean that the usability of the user interface was without imperfections. Isias and Viana [11] mention the importance of user-friendly visual representations for optimal interpretation of the collected data. Therefore, there may be room for improvement, for instance by showing the student graph while hovering over the name, without having to click on it. Some participants were confused about the difference between progress by attempt and progress over time. Many of the participants found it confusing how some students had more attempts than others. This may be because the participants are not familiar enough with how the SmartU student view is functioning. Previous studies states that it should be easy for a lecturer to understand a visualization, to avoid confusion and frustration [4, 11, 12]. For new users of the system, it could be useful to see an optional tutorial or explanation of how the student view works.

A strong positive correlation between the perceived usefulness of the learning progress graph and the perceived usefulness of the overview graphs was found. The degree of perceived usefulness across the learning progress graph and the overview graphs could therefore be connected, and it may imply that the teachers or TAs using these features, use them together. On the other hand, since many did not look at the two other tabs in the overview chart, the results may have been distorted as they had to select the neutral option in the questionnaire. Bao et al. [20] suggest that providing alternate visualizations from the same data better supports teacher's facilitation approaches. The two graphs both visualize the status of the class scores, with the overview summarizing the results of all students in the activity, and the learning progress graph showing the status of individuals.

The overview visualization received positive feedback on the colors and how uncomplicated

and easy to understand the doughnut chart is. This is in accordance with the findings from previous studies about color coding and getting information from a brief glance on the dashboard [10, 19, 22]. From the observations, we noticed that very few of the participants looked at the other tabs in the overview card, resulting in only looking at the doughnut card. If the reason for this is shortage of time, not noticing the tab, or that they just did not care for the other tabs is difficult to say. Those who did look at the other tabs found them a bit confusing and less intuitive compared to the doughnut chart. A reason for this could be, again, because of the lack of knowledge about the student view.

The mastery grid got very mixed feedback compared to the other visualizations. During the user tests we observed that some participants used the mastery grid to identify difficult topics and that they appreciated the color coding. Some participants found the sorting confusing and could not understand how a student with more red topics could be at a lower risk than the student over with less red topics. An interviewed teacher stated that they felt the information displayed on the mastery grid was inconsistent compared to the learning progress graph. A possible reason for both comments is the Simpson's paradox [38]. The Simpson's paradox is a phenomenon in probability and statistics that appears when there are several groups with different data size, but can wrongly be interpreted as the same size. Since the quizzes are adaptive, and the number of questions and topic distribution differ from time to time, the statistics can be misleading. For instance, one student can have one questions in one topic, two questions in another, and eight of the questions in a third topic. This can for example lead to 0% correctness in the first topic, 50% in the second, and 75% in the third. This can make the teacher believe the student is struggling in the first and second topic, but in reality, the student only answered 1 question incorrectly in both topics, which is far from enough to evaluate a student's performance in a topic. To make sure the mastery grid is interpreted correctly and supporting the teachers in reaching their objectives, it could be more appropriate to show a fraction of the correct answers vs. total answers instead of a percentage.

Among other feedback for the mastery grid was suggestions for other features such as displaying the average topic score for the class and total count of red students in a topic. This is in line with findings in previous studies that overall class mastery of concepts in addition to individual mastery is something teachers pay attention to [6, 13, 16]. Also, a feature for comparing two or more students was suggested by a TA. This is consistent with findings from previous studies as well [11]. Another TA wanted to see what questions the student answered wrong when clicking on the student in the mastery grid. All three of these suggestions are in line with Isaias and Viana's [11] three categories for useful information on teacher-facing dashboards: descriptive data, comparative data, and prescriptive data. Descriptive data is referring to an overview of the course's progress and student mastery, comparative data is referring to compare information between students or classes from previous years and prescriptive data is historical data, such as previous answers or data from previous years, to understand why students performed as they did, and possibly change the lecture plan. These are all features that could be implemented to further support the teachers' objectives.

To help the teachers evaluate the information and reach their objectives even faster and

more efficiently, we have implemented highlights and notifications. Previous studies have shown that this is a missing feature in teacher-facing LADs [10, 5, 11, 12, 14]. In SmartU they are triggered by a set threshold and will disappear when the user clicks on the highlighted card. The only exception is a notification beside the names of the students at risk in the mastery grid, as this notification will only disappear when the student is doing better on their attempts and not at risk anymore. From the results we found that most of the participants found the notifications to be helpful, and other were more neutral and did not pay attention to them. This is consistent with previous studies, where most teachers found notifications helpful, but a few people did find them rather disruptive [10, 11, 12]. Some participants mentioned an option to turn the notifications on and off would be optimal. Other features such as the size of the notification and different icons and colors on the notification is also something that was mentioned on the interviews, and should be considered to help teachers identify problems or identify parts of the curriculum that are going well.

6.2.2 Perceptions towards content management

Previous research has found that teacher's enjoy the ability to manage content such as editing multiple-choice quiz questions [13]. A lot of comments on the content management features of the application were related to the user experience, with some discontent regarding unintuitive buttons, clickable areas and menu options. One example was the question list in which questions could be edited by double-clicking on them. This in particular was an area of the user test where users struggled, and in turn a large portion of the feedback ended up relating to the user experience. However, there were also expressions of excitement upon discovery that questions could be edited directly from the highlighted statistics. The ability to upload questions from a file was well received, although this feature also caused some crashes during the user test which could have prevented participants from giving useful feedback. Many users were interested in using the highlighted statistics to identify and change questions when asked to take action, which indicates that this is a feature that could be very useful. The weaknesses found in the design which hindered the usage of the content management features should be improved upon and re-evaluated.

6.2.3 Perceptions towards customization

The objectives and work tasks differ for teachers and TAs, as well as between teachers at different levels of education. Differences between universities, institutes, faculties, and courses could change the requirements that a user would expect from the LAD. Changing of threshold values, resizing and moving elements in the dashboard, and allowing to set time ranges to display data from were suggestions coming from the interviews. Suggestions on visualizations showed that some would also like aggregated data rather than data on individual students. These are strong indicators that customization is important not only for usage of the dashboard at different levels of education, but for usage in different classes within the same education level.

Going back to participants' perception on using the dashboard for intervention, we saw that teachers in higher education were skeptical of using the system to contact individuals, while TA's and teachers at lower levels of education were more positively geared towards one-on-one interventions. From a customization perspective, this clearly shows that teacher's objectives or methods to reach those objectives are not one-size-fits-all, further suggesting that the system must be tailored to fit individual users' needs.

In relation to privacy concerns regarding the display of individual student scores, customization could also allow schools to adopt the dashboard in accordance with their own privacy guidelines.

6.3 Attitude towards an LAD

The third research question we posed in this thesis was: "What are teachers' motivation and attitudes towards using a LAD?" The interviews gave some clear indicators on how different teachers and TAs felt about using the system. The data collected from the questionnaire has too low sample size to differentiate between teachers at different levels of education, but gives an overall impression of motivation and attitude. Through the qualitative and quantitative data collected we have found different attitudes among teachers and TAs.

While there were no outright negative responses towards attitude or motivation, some voices were more critical and could not see themselves using the system. All TAs claimed they would use the application. Teachers were mostly positive, with only 1 out of 4 claiming they would not use SmartU as it did not fit with their style of teaching. A second teacher was positive, but claimed they would only use the system once in a while. As discussed in 6.2.3 the system should be customizable if it intends to be widely adopted, as the requirements for each class or course vary independent of education level. Regardless of the system introducing customizability to the teacher's view, it is still limited by the student application only containing adaptive quizzes. If the learning material – in this case adaptive quizzes – offered to the students do not fit with the teaching style of the course it does not matter if the teacher's view offers tailored data visualization or is easy to use.

6.4 Implications

Our study evaluates a teacher-facing dashboard across teachers from multiple levels of education, including teaching assistants at universities. Previous studies have been limited to a specific educational level, where the implications may not be applicable to support teachers at other levels. This study verifies findings from the literature, such as the usefulness of various features, including highlighting and notifications, topic mastery, learning progress and statistics showing potential errors or misconceptions in the learning material. At the same time the nature of our study, involving teachers at multiple different education levels, reveals that findings and implications from previous research is not always

applicable and that they may be context specific.

One of the main technical implications for future dashboard design from our findings is the need for customizability. Customizability is required not just because teachers at different levels have different requirements for a dashboard, but because it also varies on a per subject basis. Customizability involves but is not limited to: selecting which visualizations to display and setting threshold values for when to trigger notifications. Other implications from our study is the need for more types of visualizations, including student comparison and class information, the addition of learning goals, options for more aggregated data on students, as well as an onboarding tutorial for first time use to prevent common misunderstandings. The additions also underpin the customization implication, as dashboard users can choose which of these features to display.

6.5 Limitations

While most of the study went according to plan, we still encountered some challenges which we will go through in this section. Because of inexperience in planning and conducting studies, we made a few mistakes while creating the questionnaire. The questions in the questionnaire could have been worded differently to avoid possible misinterpretations. The questions should also have an option for not answering a question if they did not notice or interact with the visualization/feature in question. Because of these mistakes, the collected data from the questionnaire was difficult to interpret. On questions about visualizations or features the participants did not interact with, they were prompted to answer the neutral option, which made the results of the analysis less dependable.

The sample size and the participants' backgrounds are also weaknesses of this research. While we had a goal to reach 20-30 participants, we only got 13 people to participate in our study. Having a sample size less than 30 makes statistical analysis not reliable [7]. Unfortunately, it turned out to be more difficult to recruit teachers and lecturers to participate in the study than we assumed. While inviting almost 100 teachers and lecturers in addition to posting the invitation on NTNU's intranet, we only got 5 responses. To get more participants we turned to TAs and managed to recruit 8 additional participants to the study. The low teacher percentage among the participants, combined with the low sample size is a shortcoming for the study. Response rates of only 10% is not uncommon [7], but in our case we got approximately a 5% response rate among the teachers.

7 Conclusions and future work

In this thesis we formed the hypothesis that an LAD can effectively support teacher intervention and adaption of learning format. To help test this hypothesis we asked three research questions to evaluate how an LAD can support these two teaching objectives, and how the motivations and attitudes differ between teachers.

In trying to answer RQ1, we found that teachers can use the dashboard to identify and contact students at risk, be it either one-on-one or on a class level. However, this approach to intervention is not necessarily the best fit for all teachers. At universities teachers may have hundreds of students and it would not be feasible to contact individuals. While the features intended to support intervention in SmartU were overall well received, the usage context of the system must be considered when determining how effectively it can support teacher intervention.

To answer RQ2, we have looked at the different visualizations individually as well as the system as a whole. The results of the study showed us that the visualization for highlighted statistics can help teachers manage their content and identify bottle necks that can help them adapt their learning format. The learning progress graph can easily help the teacher see the class average and how many students are struggling, giving an indication of how the current lecture plan is received. Similarly, the overview also gives an indication of how the class is doing. The mastery grid visualization helps investigating individual students and helps the teacher identify topics students are struggling with, which can help them adapt parts of the curriculum or lesson plan.

Managing and sorting content was also well received by the participants in our study, even though there are room for improvement related to the user experience. Being able to see statistics for each question and sorting by topic or performance seemed to be helpful for the teachers' evaluation of the quality of questions and lectures. Notifications can help teachers identify problems quicker, and thus save time. On the other hand, depending on the class size and what level the teacher is teaching at, it could be more appropriate to show more data on class level, and less data on individual students. Since the system only supports quiz data, it is not suitable for all subjects either. In order for the LAD to have better use at different levels of education for adapting teaching, lesson plan or learning format, we have seen strong indications from the results that customization is important.

Finally, answering RQ3 we found that there are different motivations and attitudes between teachers towards the LAD. Teaching assistants and those teaching at primary to upper secondary school were more positive about using the system as a tool in their work. On the other hand, those teaching at universities who may have large number of students in

their courses are less interested in the features which facilitates intervention on individuals or changing individual quiz questions. The data being displayed originating from a strictly quiz-based system also did not fit with the teaching style of some teachers. In short, while positively received it does not fit in all contexts which does affect the motivations and attitudes of teachers towards the SmartU dashboard.

7.1 Future work

Based on the findings and limitations, we have a few suggestions for future research and further development of the artifact. Our suggestions can be summarized as follows:

- Carry out a new experiment with a larger selection of participants with teaching experience from different teaching levels. Create better questionnaires for better quantitative data analysis and let the user tests be longer, possibly over a longer period of time.
- Fix UX issues related to content management. This can be done by having more design iterations with user tests (2-3 users) between each iteration to see if the changes were beneficial.
- Add new features that have been suggested by the participants in this study. These are features like:
 - Customization: Choosing, changing and moving elements in the dashboard, being able to change thresholds for risks.
 - Compare student and class information
 - Adding learning goals
 - Being able to choose between individual student data and aggregated data on students.
 - Onboarding tutorial to avoid the common misunderstandings about the different visualizations
- The artifact is just a prototype and is not ready to be deployed as it is. There are no login/user account features, no authorization and the database connection pool is always open while the application is running. In addition, the backend of the current SmartU student view needs to be updated to support questions with topics. These are things that must be taken into account when further developing the application.

Bibliography

- [1] Zacharoula Papamitsiou et al. ‘Supporting learners in a crisis context with smart self-assessment’. In: *Radical Solutions for Education in a Crisis Context*. Springer, 2021, pp. 207–224.
- [2] Phillip Long. *Proceedings of the 1st international conference on learning analytics and knowledge*. ACM, 2011.
- [3] Michail Giannakos. *Experimental Studies and Methods in Child-Computer Interaction and Learning Technology Research*. 2021.
- [4] Juan Zheng et al. ‘Self-regulation and emotion matter: A case study of instructor interactions with a learning analytics dashboard’. In: *Computers & Education* 161 (2021).
- [5] Christothea Herodotou et al. ‘Implementing predictive learning analytics on a large scale: the teacher’s perspective’. In: *Proceedings of the seventh international learning analytics & knowledge conference*. 2017, pp. 267–271.
- [6] Franceska Xhakaj, Vincent Alevén and Bruce M McLaren. ‘How teachers use data to help students learn: Contextual inquiry for the design of a dashboard’. In: *European Conference on Technology Enhanced Learning*. Springer. 2016, pp. 340–354.
- [7] Briony J Oates. *Researching Information Systems and Computing*. Sage, 2005. ISBN: 9781412902243.
- [8] Philipp Mayring et al. ‘Qualitative content analysis’. In: *A companion to qualitative research* 1.2 (2004), pp. 159–176.
- [9] John Brooke. ‘SUS: a retrospective’. In: *Journal of usability studies* 8.2 (2013), pp. 29–40.
- [10] Christothea Herodotou et al. ‘A large-scale implementation of predictive learning analytics in higher education: the teachers’ role and perspective’. In: *Educational Technology Research and Development* 67.5 (2019), pp. 1273–1306.
- [11] Pedro Isaias and Adriana Backx Noronha Viana. ‘On the Design of a Teachers’ Dashboard: Requirements and Insights’. In: *Learning and Collaboration Technologies. Designing, Developing and Deploying Learning Experiences*. 2020, pp. 255–269.
- [12] Konstantinos Michos et al. ‘Involving teachers in learning analytics design: Lessons learned from two case studies’. In: *Proceedings of the Tenth international conference on learning analytics & knowledge*. 2020, pp. 94–99.

- [13] Neil T Heffernan and Cristina Lindquist Heffernan. ‘The ASSISTments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching’. In: *International Journal of Artificial Intelligence in Education* 24.4 (2014), pp. 470–497.
- [14] Inge Molenaar and Carolien Knoop-van Campen. ‘Teacher dashboards in practice: Usage and impact’. In: *European conference on technology enhanced learning*. Springer. 2017, pp. 125–138.
- [15] Vincent Aleven et al. ‘Developing a teacher dashboard for use with intelligent tutoring systems’. In: *technology* 34 (2010), pp. 44–50.
- [16] Frančeska Khakaj, Vincent Aleven and Bruce M McLaren. ‘Effects of a teacher dashboard for an intelligent tutoring system on teacher knowledge, lesson planning, lessons and student learning’. In: *European conference on technology enhanced learning*. Springer. 2017, pp. 315–329.
- [17] Sara De Freitas et al. ‘Foundations of dynamic learning analytics: Using university student data to increase retention’. In: *British journal of educational technology* 46.6 (2015), pp. 1175–1188.
- [18] Antonette Shibani, Simon Knight and Simon Buckingham Shum. ‘Educator perspectives on learning analytics in classroom practice’. In: *The Internet and Higher Education* 46 (2020), p. 100730.
- [19] Mingyu Feng, Neil Heffernan and Kenneth Koedinger. ‘Student modeling in an intelligent tutoring system’. In: *Intelligent tutoring systems in e-learning environments: Design, implementation and evaluation* (2011), pp. 208–236.
- [20] Haogang Bao et al. ‘The effects of a learning analytics dashboard on teachers’ diagnosis and intervention in computer-supported collaborative learning’. In: *Technology, pedagogy and education* 30 (2021), pp. 287–303.
- [21] Raphael A Dourado et al. ‘A Teacher-facing Learning Analytics Dashboard for Process-oriented Feedback in Online Learning’. In: *LAK21: 11th International Learning Analytics and Knowledge Conference*. 2021, pp. 482–489.
- [22] Yeonjeong Park and Il-Hyun Jo. ‘Factors that affect the success of learning analytics dashboards’. In: *Educational technology research and development* 67 (2019), pp. 1547–1571.
- [23] Charles Lang et al. *Handbook of learning analytics*. SOLAR, Society for Learning Analytics and Research New York, NY, USA, 2017.
- [24] Figma. *About Figma: the collaborative interface design tool*. URL: <https://www.figma.com/about/> (visited on 09/05/2022).
- [25] IntelliJ IDEA. *Why IntelliJ IDEA*. URL: <https://www.jetbrains.com/idea/> (visited on 09/05/2022).
- [26] Visual Studio Code. *Documentation for Visual Studio Code*. URL: <https://code.visualstudio.com/docs> (visited on 09/05/2022).
- [27] GitLab. *About GitLab*. URL: <https://about.gitlab.com/company/> (visited on 09/05/2022).
- [28] Docker. *About Docker*. URL: <https://www.docker.com/company/> (visited on 09/05/2022).

- [29] React. *Getting Started - React*. URL: <https://reactjs.org/docs/getting-started.html> (visited on 09/05/2022).
- [30] TypeScript. *TypeScript: JavaScript With Syntax For Types*. URL: <https://www.typescriptlang.org/> (visited on 09/05/2022).
- [31] Node.js. *About Node.js*. URL: <https://nodejs.org/en/about/> (visited on 09/05/2022).
- [32] Express. *Fast, unopinionated, minimalist web framework for Node.js*. URL: <https://expressjs.com/> (visited on 09/05/2022).
- [33] MySQL. *MySQL*. URL: <https://www.mysql.com/> (visited on 09/05/2022).
- [34] Jakob Nielsen. ‘Iterative user-interface design’. In: *Computer* 26.11 (1993), pp. 32–41.
- [35] Jakob Westermoen and Martin Lunde. ‘SmartU: Investigating the effects of visualizations in adaptive self-assessment systems’. MA thesis. NTNU, 2020.
- [36] John Brooke. ‘System usability scale (SUS): a quick-and-dirty method of system evaluation user information’. In: *Reading, UK: Digital equipment co ltd* 43 (1986), pp. 1–7.
- [37] Fred D Davis. ‘Perceived usefulness, perceived ease of use, and user acceptance of information technology’. In: *MIS quarterly* (1989), pp. 319–340.
- [38] Clifford H Wagner. ‘Simpson’s paradox in real life’. In: *The American Statistician* 36.1 (1982), pp. 46–48.

Appendices

A Study

A.1 Form of consent

Are you interested in taking part in the research project "User test of Learning Analytics Dashboard for teachers"?

This is an inquiry about participation in a research project where the main purpose is to evaluate elements in a teacher-facing dashboard. In this letter we will give you information about the purpose of the project and what your participation will involve.

Purpose of the project

The learning analytics dashboard (LAD) SmartU for teachers is the subject of a master thesis where the aim is to investigate how elements such as visualizations of student data and other features support teachers' objectives, such as identifying students at risk, intervention, and adapting the learning format or strategy in a university course. A prototype of the LAD has been developed and will be the artifact of the study, where the aim is to measure the effect of the elements in the LAD with regards to supporting teachers. Teachers will be the subjects of the study, where they will get a firsthand experience of using the new system. An Associate Professor at the Department of Computer Science at NTNU Trondheim will be responsible for the project.

What does participation involve for you?

Participation in the study includes a user test, followed by an interview and a short questionnaire, which in total is expected to last approximately 45 minutes.

The user test includes hands-on testing of the SmartU prototype, where we will make notes from our observations of your interaction with the system. These notes will include how you navigate the system, how you make sense of the different elements, verbal comments from you while you use the system etc.

After the user test there will be a short interview asking you about the experience of using the system. Questions will include topics such as how you felt about the SmartU system, previous experiences with similar systems etc. Then there will be a questionnaire to gather some quantifiable measurements.

The data collected will include some information about you such as name, age, gender, how long you have been in your profession, what level of education you teach at.

Participation is voluntary

Participation in the project is voluntary. If you choose to participate, you can withdraw your consent at any time without giving a reason. All information about you will then be made anonymous. There will be no negative consequences for you if you choose not to participate or later decide to withdraw.

Your personal privacy – how we will store and use your personal data

All personal data will be treated confidentially. Only the project group (see general information section below) will have access to the personal data. The list of names of the participants will be stored in NTNU Sharepoint according to the data processing agreement between NTNU and Microsoft. Only the researchers and data controller will have access to this site.

Questionnaires will be hosted by UiO Nettskjema.

What will happen to your personal data at the end of the research project?

We state that the participants will not be recognizable in the publication. The project is scheduled for completion by June 2022, then all data will be anonymized.

Your rights

So long as you can be identified in the collected data, you have the right to:

- access the personal data that is being processed about you
- request that your personal data is deleted
- request that incorrect personal data about you is corrected/rectified
- receive a copy of your personal data (data portability), and
- send a complaint to the Data Protection Officer or The Norwegian Data Protection Authority regarding the processing of your personal data

What gives us the right to process your personal data?

We will process your personal data based on your consent.

Based on an agreement with NTNU, NSD – The Norwegian Centre for Research Data AS has assessed that the processing of personal data in this project is in accordance with data protection legislation.

General information - project group:

The supervisor of the project is Michail Giannakos, Associate Professor at Department of Computer and Information Science at NTNU, e-mail: michailg@ntnu.no, address: Sem Sælands vei 9, IT-bygget 103, phone number: +47 73590731.

Co-supervisor of the project is Zacharoula Papamitsiou, Research scientist at SINTEF Digital Trondheim, e-mail: papamitsiou.xara@gmail.com, address: S P Andersens vei 3

If you would like to participate or have any questions concerning the project, please contact:

Andine Luick, e-mail: andinel@stud.ntnu.no,
Master student at the Department of Computer Science (IDI) at NTNU

or

Fredrik Monsen, e-mail: fredrmon@stud.ntnu.no,
Master student at the Department of Computer Science (IDI) at NTNU

Data Protection Officer (Personvernombud) at NTNU (Thomas Helgesen, thomas.helgesen@ntnu.no)

The study has been notified to the NSD – The Norwegian Centre for Research Data AS (personverntjenester@nsd.no, 55 58 21 17) has assessed that the processing of personal data in this project is in accordance with data protection legislation.

Yours sincerely,

Project Leader

(Researcher/supervisor)

Student (if applicable)

Consent form

I have received and understood information about the project and have been given the opportunity to ask questions. I give consent:

- to participate in *user test with written observations*
- to participate in *an interview*
- to participate in *an online survey*

I give consent for my personal data to be processed until the end date of the project, approx. *6th of June 2022*

(Signed by participant, date)

A.2 Observation guide

Observation guide

Introduction: The user test will begin with an introduction from the observers, who they are, the purpose of the user test, and some information of how the test will be conducted. The tone of the introduction should help establish a relaxed environment between the observer(s) and test-subject.

Interaction with test-subject: The observer should not interact with the test-subject, except for when the test-subject asks for help. It is important to not spend too much time on helping the test-subject.

Capturing contextual information: Information about room ambience and atmosphere should be written down. Observations will also include notes on how the test subject's emotional state is during the test, and whether or not it changes over time.

Capturing crucial information: During the test, notes should be taken while observing the test-subject. The notes should summarize the events and unexpected actions the test-subject made.

Examples of events/actions to be noted:

Category	Includes	Researchers should note
Test environment	Set up, ambience, atmosphere	Differences between environment if there are changes between subjects
Emotional state	The emotional state of the test subject (stressed/happy/angry etc.)	Changes in emotional state, any particular aspect of the artifact that probes a specific reaction, is the subject focused and/or motivated
Unexpected events	System bugs/crashes	If anything unexpected occurs note when, how and why
Help	Questions, frequency	What questions does the test-subject ask, and how often do they ask for help
Disturbances	Noises, people entering/leaving the room, phone ringing, etc.	Unwanted noises that can affect the test-subject's concentration and focus on the system

End of testing and observations: Thank the test-subjects for their time and provide them with the observer's contact details.

A.3 Interview guide

Interview Guide

Capturing answers: Answers will be captured through written notes and recording of audio. This ensures that no information is lost and can be retrieved if the notes are insufficient.

Develop a rapport with the respondent:

Ask questions that lead detailed answers: Questions will be open-ended rather than close-ended, and posed in such a manner that the interview subject provides detailed answers. The interviewer will have a set of questions to go through, but may ask follow up questions based on the answers given to either clarify or expand on the answer. Questions will be asked one by one.

Examples of questions:

- How did you feel about using the SmartU dashboard?
- What was your most preferred visualization in the dashboard and why?
- What was your least preferred visualization in the dashboard and why?
- How do you think you could utilize the information presented in the dashboard for something useful?

End the interview: When all questions are exhausted, summarize the key points that the interviewer feels the respondent has provided. This gives the respondent a final chance to expand or clarify any points. Thank the respondents for their time and provide them with the interviewer's contact information.

A.4 Questionnaire

SmartU

Side 1

Obligatoriske felter er merket med stjerne *

Personalialia

Name (first name and last name) *

Age *

Gender *

Email *


I have worked as a... *

Teacher/lecturer

Teaching assistant


Other

Teaching experience *

 Dette elementet vises kun dersom alternativet «Teacher/lecturer» er valgt i spørsmålet «I have worked as a...»


How many years you have worked as a teacher/lecturer

Teaching experience *


 Dette elementet vises kun dersom alternativet «Teaching assistant» er valgt i spørsmålet «I have worked as a...»

How many semesters you have worked as a teaching assistant

What other teaching experience do you have? *

 Dette elementet vises kun dersom alternativet «Other» er valgt i spørsmålet «I have worked as a...»

Teaching level *

 Dette elementet vises kun dersom alternativet «Teacher/lecturer» er valgt i spørsmålet «I have worked as a...»

At what level do you teach? If you have taught at multiple levels, select the highest

Primary school

Middle school

Upper secondary / high school

College

University

Obligatoriske felter er merket med stjerne *

Previous experience with learning technology

Do you have any previous experience using Learning Analytics tools? *

Yes

No

Which Learning Analytics tools or applications do you have experience with? *

i Dette elementet vises kun dersom alternativet «Yes» er valgt i spørsmålet «Do you have any previous experience using Learning Analytics tools?»

E.g. Blackboard Analytics for Learn

i Dette elementet vises kun dersom alternativet «Yes» er valgt i spørsmålet «Do you have any previous experience using Learning Analytics tools?»

Frequency of Learning Analytics usage

Very often Often Sometimes Rarely Very rarely

How often do you use Learning Analytics software? *

Obligatoriske felter er merket med stjerne *

Usability

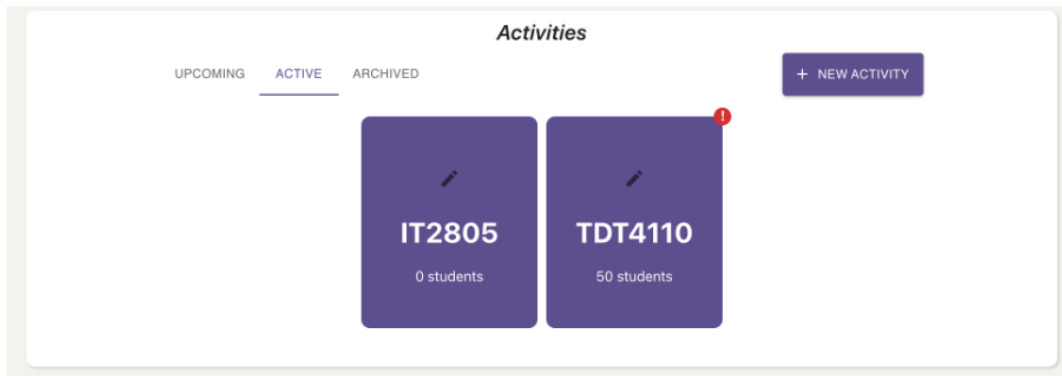
System Usability Scale

Select an option between 1-5, where 1 is Strongly disagree and 5 is Strongly agree

	1	2	3	4	5
I think that I would like to use this system frequently	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found the system unnecessarily complex	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I thought the system was easy to use	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I think that I would need the support of a technical person to be able to use this system	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found the various functions in this system were well integrated	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I thought there was too much inconsistency in this system	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I would imagine that most people would learn to use this system very quickly	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found the system very cumbersome to use	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I felt very confident using the system	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I needed to learn a lot of things before I could get going with this system	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

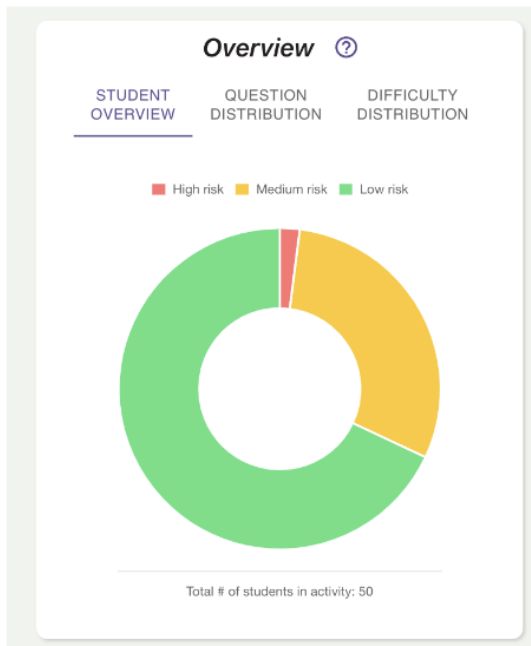
Evaluation of the usability of SmartU

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
It was easy to change a question in the editor *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It was easy to understand which questions needed changing *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It was easy to add new questions to an activity *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It was easy to add new announcements *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It was easy to make sense of the color-coding in graphs and visualizations *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It was easy to understand that the visualizations showed the learning progress of the students *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I could make sense of the mastery grid to see which topics are easy *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I could make sense of the mastery grid to see which topics are difficult *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I could make sense of the learning progress graph *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It was easy to understand the sorting of the student list for the learning progress visualization *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It was easy to understand the sorting of the student list for the mastery grid *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It was easy to create a new activity *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



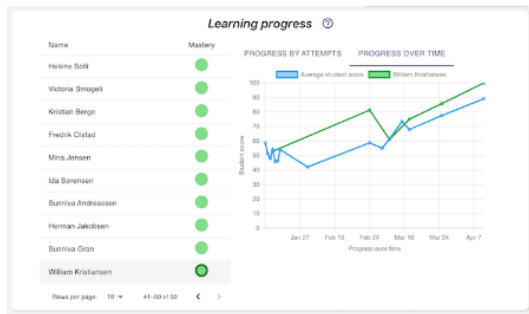
Usefulness of activity selection

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I could easily understand in which activity there was need for extra attention *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is useful to have activities split across three tabs (upcoming, active and archived) *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Usefulness of overview

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I found it useful to see an overview of the students' potential risk of failing in a doughnut chart *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found it useful to see how the questions is distributed over the different topics *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Knowing the topic distribution can be a useful tool to have when adding new questions *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found it useful to see how the questions is distributed over the different difficulties *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Knowing the difficulty distribution can be a useful tool to have when adding new questions *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Usefulness of learning progress visualization

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I believe I could use the Learning Progress graph to check if the students are on track *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found it useful to be able to switch between progress by attempts and over time *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Mastery grid ⓘ

Name	CSS	HTML	JavaScript
▲ Tuva Fjeld	30	40	35
▲ William Wold	24	47	36
Kristoffer Aalerud	43	33	52
Thea Christiansen	47	54	34
Celine Fossum	47	23	67
Hedda Nilsen	42	52	49
Maren Reed	53	44	48
Noah Strand	47	52	47
Mathias Vegge	61	42	45
Karoline Nordskaug	58	56	38

Rows per page: 10 ▾ 1–10 of 50 < >

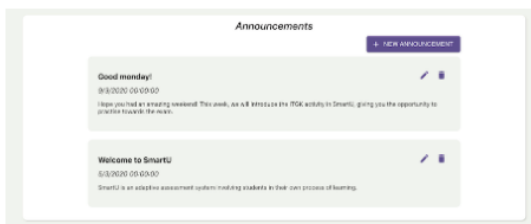
Usefulness of mastery grid

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I believe the Mastery Grid could help me identify students at risk of failing *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found it useful to be able to click on a student's name to get more details about the student *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Usefulness of activity insights

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I found the question statistics to be a useful tool to identify questions with potential issues *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found the question statistics to be a useful tool to identify questions students find difficult *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found the question statistics to be a useful tool to identify questions students find easy *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found the login statistics to be useful to see recent student activity *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Usefulness of announcements

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I found the announcements to be useful to reach out to students with messages *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

B Interviews

Interview participant 01

I: Først, hva synes du egentlig om å bruke SmartU dashboardet?

P01: Det var bra. Det var veldig lett å sette seg inn i. Det var ikke for mye. Jeg følte ikke det var noe mer enn det trengte å være. Det var oversiktlig. Veldig pent. Ja, det var en glede.

I: Kan du nevne en eller flere funksjonaliteter som du synes var ekstra nyttig, og hvorfor?

P01: Hmm... Denne statistikken da, som highlighter hvilke spørsmål som var vanskelig. Eller spørsmål som folk bruker ganske lang tid på. Det var jo ganske nyttig, for det gir jo... kan gi tilbakemelding om man har skrevet spørsmålene på en dårlig måte, eller rett og slett for vanskelig. Mmm, den læringsprogresjon tror jeg var veldig nyttig, for hver enkelt elev og overall. For da har man jo, får man jo ganske, sånn... eller det kan være en god indikasjon på om undervisninga du holder og opplegget du har er greit eller ikke.

I: Synes du det var noe funksjonalitet som var forvirrende eller vanskelig å forstå?

P01: Det var vel den hiccupen med å lage et nytt fag, men selv prosessen var jo bra.

I: Jeg så at du prøvde å registrere uten å skrive en description først. Var det noe i forhold til forms der som var forvirrende, foruten om at det ikke funka?

P01: Nei jeg bare så ikke asterisken. Jo, de announcements-greiene, jeg visste ikke om det var for alle fagene jeg hadde eller for ett og ett fag, så det var jeg ikke...

I: Ja så du ble litt usikker på hvor det hadde blitt publisert hvis du hadde skrevet en?

P01: Ja.

I: Nå noe spesifikt om visualiseringer, der du har grafene, donut-grafer og progresjonsgrafene, og noen flere grafer som du ikke kikket på. Så da er spørsmålet, hvilke av visualiseringene likte du best, og hvorfor?

P01: Mmm, nei det var vel progresjon da. Den tror jeg var lettest å forstå og...eh...vet ikke om jeg likte den så sykt mye bedre enn andre grafer, men ja den var veldig lett å forstå. Alle grafene – eller de jeg så på – var lette å forstå.

I: Den progresjonsgrafene hadde jo to moduser, én der hvor man kunne se for hver attempt og én for over tid. Jeg så at du bytta mellom dem, var det forståelig hva forskjellen på de var?

P01: Hmm, vet ikke om jeg tenkte på det egentlig. Vet ikke... jeg tror jeg tenkte at den første grafen jeg så var bare for alle gjennomsnittlig, så når jeg trykte på en person så tenkte jeg at det var individuelt.

I: Motparten til det spørsmålet: hvilke av visualiseringene likte du minst, og hvorfor? Var det noen som utpekte seg som "den verste"?

P01: Donut-grafene, for den kunne fint vært en pie-diagram ha-ha-ha.

I: Var det noe funksjonalitet du skulle ønsket var tilstede på dashbordet, kanskje noe som du hadde forventet?

P01: Jeg prøvde jo å trykke meg inn på en person for å se om jeg kunne sende en direkte melding på den plattform, men det er kanskje ikke meningen. Ehh, ja det var vel egentlig bare det at du kunne gi direkte tilbakemelding til den personen.

I: Tror du at du kunne brukt den informasjonen som ble presentert på dashbordet til noe nyttig, i den arbeidssettingen som student assistent eller lærer?

P01: Ja. Jeg vet ikke som lærer, eller som lærer kan du jo se om det er noe av opplegget ditt du må forbedre da. Og som stud.ass. kan det jo være at du kanskje prøver å nå ut til de du veit – eller prøver å gjøre deg ekstra tilgjengelig for de personene som du ser kanskje sliter litt i faget.

I: Du snakka jo litt om det egentlig, ja du har egentlig besvart det neste spørsmålet mitt, med hvordan du tror dashbordet kunne hjulpet med å nå ut til studenter som står i fare for å stryke; du ville jo sende direkte melding. Du prøvde jo også å skrive en announcement. Så det spørsmålet trenger vi kanskje ikke å gå gjennom på nytt. Hvordan tror du at – og dette har du forsåvidt også svart på men vi kan ta det kjapt – hvordan tror du dashbordet kunne hjulpet deg med å gjøre passende endringer til læreplanen, hvis du antar at du har rettighetene til å gjøre det?

P01: Ja det er jo å se på metrikken som tilsier at det er ting folk sliter med.

I: Du nevnte jo at du synes det var sånn akkurat passe når du gikk inn. Du følte ikke at det var enten for simplistisk eller for avansert?

P01: Nei.

I: Så var det disse notifikasjonene, som var i dashbordet, disse røde dottene. Følte du at de hjalp deg med å fokusere?

P01: Ja det vil jeg absolutt si.

I: Så du føler at det hjelper deg med å finne det du...

P01: Det som trenger oppmerksomheten min først ja. Det vil jeg absolutt si.

I: Så du tror du ville foretrukket å hatt det versus å ikke ha det?

P01: Ja, hvis jeg prøver å være en god læringsassistent så vil jeg nok tenkt at det er noe nyttig.

I: I et system som dette her så gjør studentene quizer over tid, så de forbedrer jo da scoren sin. Spørsmålet da er: hvordan score foretrekker du å se, er det fra siste forsøk eller er f.eks. average av totalen. Gjorde du deg noen tanker om det mens du holdt på?

P01: Jeg var ikke klar over at det var siste forsøk og første forsøk, så det vet jeg jo egentlig ikke hva jeg tenkte om da. Men jeg tenker jo at det kanskje er mest nyttig å se på siste forsøk, for det vil jo være det som er representativt for det personen kan akkurat der og da.

I: Systemet som de svarer på spørsmålene i er jo et adaptivt system, så det er ikke nødvendigvis sånn at de får like mange spørsmål innenfor hver topic. Så hvis de da hadde fått ett spørsmål fra en av topicene og svart feil på det, men kanskje hadde vært god på det ellers, hva føler du om det i forhold til hvilken score som skal vises frem da?

P01: Godt spørsmål. Ja da ville det vært mest nyttig med gjennomsnitt eller med mest representativt.

I: Alle siste spørsmål, hvor ofte tror du at du ville brukt en slik applikasjon som SmartU?

P01: Ville nok tatt en titt i uka. Hvis jeg regner med at jeg er stud.ass. en gang i uka så ville jeg nok tatt en titt før jeg skal på jobb.

Interview participant 03

I: Hva synes du om å bruke SmartU dashbordet?

P03: Dashbordet.. Mener du den første sida der alle aktivitetene står, eller er du inne på faktisk statistikken?

I: Jeg tenker på den statistikkdelen

P03: Statistikkdelen? Ja, jeg synes den var veldig praktisk og enkel og intuitiv. Sånn for min del da. Jeg synes var lett å finne frem til de tingene jeg logisk prøvde å... logisk tenkte at ok, for å bruke det her må jeg se litt på det her da. Jeg synes det var oversiktlig, jeg synes det var ganske logisk og praktisk anlagt.

I: Kan du nevne en eller flere funksjonaliteter som du likte ekstra godt og syntes var ekstra nyttig?

P03: Jeg synes grafen som viser over tid, altså utviklingen by attempt eller over tid. Også det der kakediagrammet eller hva du kaller det. Den distribusjonen var også interessant å se. Det gir deg en ide over hvor elevene dine ligger, og om det kanskje er noen elever du bør følge opp litt, eller om det er noen som synes det er for enkelt.

I: Var noen av funksjonalitetene forvirrende eller vanskelige å forstå?

P03: Hmm... egentlig ikke. Hvis jeg hadde hatt litt mer kunnskap om tema, så hadde ikke det å se på spørsmålene vært noe problem. Så jeg synes det var ganske greit.

I: Det her spørsmålet er litt likt et tidligere spørsmål, men hvilken visualisering likte du best? Hvis du klarer å se for deg en visualisering du likte best.

P03: Jeg synes kanskje... jeg personlig synes at vanlig grafer og den kakegreia var det enkleste, fordi når du kom på vanskelighetsgraden, altså søylediagrammet, kanskje var litt mindre... det var full mulig å forstå det, men det var ikke så intuitivt å se med en gang hva greia var.

I: Skjønner. Så den visualiseringen du likte minst det blir kanskje...?

P03: Ja, det blir søylediagrammet.

I: Hva det en annen funksjonalitet du skulle ønske var tilstede på dashbordet?

P03: Hmm... Jeg kommer ikke på noe. Fordi... eller... hvis du trykket på redigering, så fikk du opp ekstra detaljer på alle spørsmålene, ikke bare de som stikker seg ut. Så da er egentlig jeg fornøyd. Og det faktisk greit at den informasjonen ikke kommer opp med en gang, for da ville det blitt veldig overveldende å se på.

I: Tror du at du kunne ha brukt informasjonen som ble presentert på dashbordet til noe nyttig? For eksempel i forhold til undervisning eller vurdering?

P03: : Ja, det tror jeg absolutt jeg kunne ha gjort.

I: Tror du det også kunne ha hjulpet deg med å nå ut til studenter som står i fare for å stryke?

P03: Så lenge de går inn og tar noen forsøk over tid, så tror jeg at det absolutt ville gitt meg en god ide om hvilke studenter som kan stryke. Så ja, det tror jeg.

I: Det kan jo være litt vanskelig å nå ut til en student gjennom denne applikasjonen siden vi ikke har noen direkte kobling mellom student og lærer. Tror du det ville vært greit å nå ut til en student utenfor systemet gjennom epost eller noe i den duren?

P03: Nja, personlig ville jeg helst snakket med eleven face-to-face om dette, fordi det kan være veldig kjipt og man kan få et veldig dårlig forhold til denne type oppgaver hvis du får mail om at du har gjort det dårlig.

I: Så du synes det er helt greit at det ikke er noen kobling der lærer og student kan kommunisere gjennom systemet.

P03: Jeg synes det er helt greit at læreren kan få se resultatene, og så kan læreren få lov til å finne ut av hvordan han eller hun vil takle denne situasjonen utenfor. Det kan hende det hadde vært greit med en tilbakemeldingsfunksjon. Det kan være en ide her. Det kan være nyttig hvis elevene ikke ser meg på noen uker, en tilbakemelding som ikke alle ser, men som blir til hver enkelt elev.

I: Hva synes du om dashbordets design. Var det for mye, var det lite, akkurat passe..?

P03: Jeg synes egentlig det var en grei mengde. Det var ikke veldig overveldende og det var ikke for lite heller. Så... Det meste jeg så syntes jeg var nødvendig hvis jeg skal gjøre en vurdering av det studentene gjør på quizene.

I: Før du trykket på de activity insights'ene, som er det vi kaller disse her da (peker på activity insights på skjermen), så var det en rød dott med utropstegn i hjørnet. Den så jeg at du la merke til og kommenterte...

P03: Ja, det røde utropstegnet ga veldig tydelig tegn på hva du burde se på først iallefall.

I: Synes du det var greit med en slik notifikasjon som kunne hinte til og finne fram til hva som kan være et problem med quizen?

P03: Ja, det synes jeg var veldig nyttig, også den røde trekanten. Eventuelt kunne man også hatt et grønt utropstegn på de tingene som stikker seg ut som veldig lett også- For det er lett å se rødt, rødt, rødt, skummelt! Men så ser man kanskje ikke de spørsmålene og tingene som ikke gir noen fremgang eller utvikling da.

I: Noe jeg også fikk med meg var at du så at den masteryfargen på grafen var basert på siste attemp.

P03: Eh.. Ja, nja, jeg lurte på det, men jeg visste det ikke.

I: Åja, men ja, det er siste attemp det er basert på.

P03: Å, det er det ja.

I: Det er det. Og vi lurte veldig lenge på hva vi skulle basere scoren. Om vi skulle bruke gjennomsnittet, eller en gjennomsnitt fra de siste attemptene. Hvis man f.eks. bruker gjennomsnittet (viser filtrering av gjennomsnitt på grafen), så kan tidligere dårlige resultater forkludre til gjennomsnittsscoren. Hva tenker du rundt dette? Hvordan kunne du tenke deg å få denne scoren presentert?

P03: Med den siste attempt greia så får du jo bare den forrige, men det er egentlig greit så lenge du får se de tidligere forsøkene i grafen også. Men hvis du ikke har skjønt det da, og bare ser på fargen, så kan de gi misvisende informasjon. Jeg tror at en gjennomsnittlig score... det kan også være litt vanskeligere å tyde. Jeg har egentlig ikke noen sånn kjempegod ide, og synes det her fungerer helt OK. Men da må man være klar over at man må se på utviklingen.

I: Vi har også noen filtreringsmuligheter som du ikke la merke til under brukertesten *viser der på grafen*.. Her man kan velge mellom siste forsøk og gjennomsnittlig.

P03: Hvis det er alternativer å bytte mellom, så er ikke det dumt.

I: Da har jeg et siste spørsmål til deg! Tror du at du ville ha brukt SmartU om du hadde hatt muligheten, og hvor ofte tror du at du ville gått inn og sjekket statistikken?

P03: Hvis jeg hadde jobbet med programmering eller hatt det her som fag...

I: Hva om du ser for deg at du skulle bruke det i et engelskfag?

P03: I engelsk? Quizer kan være nyttig i engelsk, så jeg synes det er absolutt et nyttig verktøy, men jeg ville nok ikke brukt det veldig ofte, men jeg ville nok hatt det som en ting jeg bruker en gang i blant for å skape variasjon i undervisningen min da. Pluss at hvis det er ting jeg ønsker å ha oversikt over mer enn... eller enklere å ha oversikt over.

I: Ok, det var alle spørsmålene jeg hadde. Takk for at du stilte til intervju!

Interview participant 04

I: Det første spørsmålet jeg har til deg er: hva synes du om å bruke SmartU dashbordet?

P04: Jeg kan si litt om måten jeg underviser. Vi bruker veldig lite quiz, iallfall ikke som en del av karakteren. Vi bruker mye tilbakemeldinger, altså skriftlige tilbakemeldinger, så for meg ville det ikke vært veldig nyttig. Men jeg kan se at det kan være veldig nyttig. Jeg vet at i større fag er det veldig vanlig at man også bruker quiz for karaktersetting. Så jeg er sikker på at det kunne vært nyttig for noen. Og så tenker jeg at fordelene er at du fanger studenter som ikke gjør det bra. Tenker at en risiko kan være at det blir veldig tallstyrt og at studenter forsvinner. Jeg prøvde å finne ut av hvor mye informasjon jeg kan finne om studenten. Og hvor greit er det egentlig å få studenter flagget med en sånn rød farge?

I: Og nå lurer jeg på om du kan nevne en eller flere funksjonaliteter som du syntes var ekstra nyttig og hvorfor?

P04: Jeg tenkte at det var fint å se den kakegrafene, overview, som viser hvor mange som ligger godt an osv. Og så syntes jeg at den grafen... jeg brukte den en del, den der learning progress grafen, syntes den også var nyttig å se. Men jeg tenker også at det her også blir veldig avhengig av hvordan man har laget quizen. Det trenger ikke å bety at studentene er dumme, men det kan like gjerne bety at spørsmålene er dårlig formulert eller at det er misforståelser i måten ting er formulert.

I: Var noen av funksjonalitetene forvirrende eller vanskelige å forstå? (litt støy utenfra under dette spørsmålet)

P04: I forhold til disse statistikkene så tenkte jeg at det kanskje ikke var tydelig hvordan... hva som var nytten. Og den "Most commonly failed question"... mulig at de [som bruker systemet] ikke var fortrolige med spørsmålene i quizen. Jeg tror det kunne hjulpet om selve spørsmålet stod der, og ikke bare "question 7". Bare som et eksempel.

I: Ja. Slik systemet er nå skal det gå an å klikke på boksen og få opp spørsmålet.

P04: Ja, jeg var innom og klikket på de under testen. Men en forbedring kunne ha vært å vist en eller annen grafisk fremvisning. Vet ikke helt hvordan det skulle vært, men det hadde også vært en mulighet å vise spørsmålet på hovedskjermen. (05:19 -17:31) Og så tenkte jeg at det også var litt inkonsistent i forhold til progresjon. Altså det som stod i grafen og på disse attempts **peker på mastery grid**. Jeg tror det kunne ha vært forbedret. Altså at når jeg trykker på navnet til en student i mastery grid, så skal jeg kunne se noe av det samme som jeg ser i grafen da.

I: Det er faktisk de samme tallene som dukker opp i både grafen og i attempts når man trykker på en student i mastery grid. Men det kan være vanskelig å se siden i grafen er det regnet om til en prosent, mens i studentoversikten står tallene som en brøk.

P04: Ja, det kan hende at når man har lært det, så vet man at det er samme.

I: Ja, men det kan altså være vanskelig å skjønne for en ny bruker.

P04: Mhm (nikker bekreftende)

I: Ja, dette er nyttig informasjon for oss det. Det neste spørsmålet ligner litt på et tidligere spørsmål, men hvilken visualisering likte du best?

P04: Jeg er kanskje en litt mer tilhenger av slike progresjonsvisualiseringer, da de viser hva som skjer over lengre tid og hvordan studenter utvikler seg. Jeg tror at hvis du tar en sånn snapshot så er det også greit, de viser litt ting, men hvis man kan kombinere det slik dere har gjort det her med overview og learning progress så synes jeg det fungerer bra. Men det er kanskje litt mer interessant å se learning progress i bare en sånn snapshot. Så det ville jeg kanskje ha brukt litt mer. For det er jo læringsprosessen vi er interessert i. Altså hvordan studenter lærer over tid. Selvfølgelig hvis du har et slikt eksamensbasert fag så er det viktigste hva du kan på akkurat eksamensdagen, men i mange fag er også prosessen viktig.

I: Hvilken visualisering likte du minst, og hvorfor?

P04: Nei... Altså jeg tenker at statistikken kunne vært litt mer informativ.

I: At den skulle vist litt mer... ?

P04: Jeg vet ikke helt hva jeg synes er feil med den, men det kan hende at det ikke burde være fokus på spørsmål, men heller læringsmål. Som jeg sa i forhold til disse temaene, så ville det kanskje vært mer interessant å se statistikk rundt temaene. For hvis du har ett vanskelig spørsmål om algoritmer, men alle er veldig flinke i algoritmer unntatt dette ene spørsmålet. Dette kan føre til at man får feil inntrykk og tror at studentene er dårlige i algoritmer. Hvis det er på tema, så kan det også være en statistikk om f.eks. "most commonly failed topic".

I: Var det noe funksjonalitet du skulle ønske var tilstede på dashbordet?

P04: Jeg tror at det er begrenset hvis det bare er quiz da. Det er jo mange forskjellige læringsaktiviteter. Quiz er kanskje det vi... altså vi bruker det ikke for vurdering, men bruker det personlig for å sjekke underveis i forelesning for å se hvor mye man har fått med seg. Bare for å illustrere det litt mer sånn i praksis. Men jeg bruker det ikke for å se hvor mye studenter har forstått av faget, da ville jeg heller brukt andre læringsaktiviteter.

I: Det neste spørsmålet er om du tror du at du kunne ha brukt informasjonen som ble presentert på dashbordet til noe nyttig? Altså gjort noen endringer i læreplanen eller noe i den duren.

P04: Hvis jeg kunne hatt tilsvarende verktøy organisert etter læringsmål, altså at det ikke ble om spørsmål og antall spørsmål, men at jeg kunne se hvilke læringsmål som er enklest å forstå. For jeg tror ikke jeg ville hatt det pr. student, men heller generelt om faget. Og så kunne jeg gjerne ha koblet det til et verktøy slik som Mentimeter. Mentimeter har også litt av sånt [som SmartU LAD], men ikke så avansert. Jeg tror det kunne vært en vei å gå videre med det her, å koble det opp mot en sånn. Vi kaller det studentengasjement, og ikke evalueringverktøy. Men engasjementverktøy. Da kunne jeg kanskje også ha brukt Mentimeter oftere. Både for studentene selv, så kunne det vært interessant hvis dataen ikke

var basert på studentnavn, at de kunne sett hvordan de ligger an i forhold til klassen for eksempel.

I: Tror du at du hadde brukt verktøyet til å identifisere og nå ut til studenter som står i fare for å stryke?

P04: Nei, da tror jeg heller studentene hadde kontaktet meg. Jeg vet ikke om det er lov å samle så mye data om enkeltstudenter en gang, så det får dere sjekke.

I: Hehe, ja, dataen vi har brukt her er ikke ekte studentdata. Men vi må absolutt sjekke ut hvilken data vi har lov til å vise frem.

P04: Hvis målet er at dette er et verktøy som skal brukes av flere faglærere burde man være helt sikker på at datatilsynet ikke kommer.

I: Ja, det er klart. Og så over til neste spørsmål: Hva synes du om dashboardets design? Hva det overveldende, lite, akkurat passe...?

P04: Jeg synes det fungerte bra jeg. Som sagt så ble det kanskje litt mye fokus på enkeltspørsmål, så det kan hende at det kan heves opp til læringsmål, og kanskje hatt en sånn database med spørsmål med forskjellige læringsmål. Slik at man kan gjenbruke det fra år til år. Og så tenker jeg kanskje litt mer aggregert data, og ikke for hver student. Man åpner liksom en sånn pandoras boks. Man har ikke tid til å sitte med enkeltstudenter. Selv om det skulle vært lov å samle og vise slik informasjon, så vil det være urettferdig å bruke mer tid på enkeltstudenter. Så jeg ville nok heller hatt mer aggregert data på klassenivå. Det kunne kanskje ha vært koblet til bakgrunnen til studentene i større fag. Slik at man kan f.eks. se at de som kommer fra Indøk sliter med python, og det kan være nyttig informasjon.

I: Ja, det høres interessant ut! Jeg vet ikke om du la merke til det, men når du startet å bruke systemet så var det noen notifications, eller røde symboler.

P04: Ja

I: Føler du det hjalp deg med å finne fram til hva som kan være viktig å se på?

P04: Ja, jeg tror det ville vært nyttig. Spesielt hvis det skal brukes over en lengre periode, og man bare logger inn av og til.

I: En annen ting vi har hatt litt problemer med å finne den beste løsningen på er hvordan man skal sortere listen etter. Om den skal sorteres etter gjennomsnittlig score, siste forsøk, gjennomsnitt av de siste forsøkene eller om det er en annen løsning for det. Akkurat nå blir du møtt med en sortering som er basert på siste forsøk, så hvis siste forsøk er dårlig vil studenten bli rød. Vet ikke om du har noen tanker rundt det?

P04: Og da er siste attempt tilpasset studenten?

I: Ja

P04: Og hvordan er tilpasningen? Unngår man det studentene feiler, eller blir det mer fokus på det studentene feiler på?

I: Det er sånn at hvis du f.eks. sliter veldig med python, så prøver systemer å gi studentene lettere python spørsmål, og når de etterhvert blir bedre i python vil de få vanskeligere python-spørsmål. Så det er adaptivt i forhold til deres kunnskapsnivå.

P04: Ja, for det tenker jeg at hvis jeg skulle hatt sånn personalisert analytics, så er det her kjernespørsmålet. Altså hvordan.. hva er denne interaktiviteten. Jeg tenker det er sånn mer pedagogisk ting, hva hjelper studenten? Ta et steg tilbake og og begynne fra et lavere nivå eller... ja, igjen, dette er tilbake til andre læringsaktiviteter. Og man skal kanskje sitte med en flinkere student og å gå gjennom en quiz. Så det er mange ting man kan gjøre da. Så det er veldig vanskelig å si hva skal det røde skal være bygd på. Men jeg ville antagelig hatt en preferanse mot det siste, at det er de siste par forsøkene de er sortert på, fordi gjennomsnittet over alle forsøkene blir ganske meningsløst. For målet er jo at du skal lære, ikke sant.

I: Ja. For vi har med en sorteringsfunksjon der man kan sortere på gjennomsnittet, og da blir listen seende helt annerledes ut (viser fram hvordan sorteringsfunksjonen og lista blir seende ut). Her ser vi f.eks. hvordan gjennomsnittet til Tuva er veldig dårlig (Tuva er rød i lista), men hun har faktisk gjort det veldig bra på slutten. Så det gir ikke den samme indikasjonen på om hun faktisk er i fare for å stryke, så å bruke en samling av de siste forsøkene er sikkert lurt ja.

P04: Mhm, mhm. Bare for å nevne det så hadde jeg også en diskusjon med [professor], for vi har jo jobbet med noe lignende i et annet prosjekt for [en etat], for der hadde de tenkt å lage noe sånt for [etatens ansatte]. Sånn at de f.eks. kunne se [masse personlig data om enkeltpersoner], så det var masse informasjon om personer eller innbyggere da. Og det fikk vi ikke lov til fra datatilsynet. Så da måtte man gjøre det omvendt. Så i dette tilfelle ville det vært at det er studenten som får se denne informasjonen om seg selv, og ikke faglæreren. Så studenten kunne gjerne ha gått inn å sett på sin egen statistikk og så kanskje ta noen valg og finne ut av hva de vil gjøre. Så det var det vi fikk beskjed om å gjøre. Så det er det jeg antar at det er noe av det samme regelverket som gjelder her også.

I: Det er det som allerede eksisterer i studentsystemet, så de ser denne dataen om kun seg selv. Så det blir vel kanskje akkurat sånn.

P04: Ja, om seg selv, og så kan de kanskje se det aggregert mot klassen, sånn som du ser det i karakterutskriften.

I: Ja, sånn tror jeg det er. Iallfall i en graf.

P04: Men der var det sånn at [etatens]-ansatte ikke kunne ta noen beslutninger basert på den dataen. Jeg vet ikke om de kunne se den, det er jeg ikke sikker på. Men hvis en student går inn og gjør noe, eller en faglærer går inn og gjør noe basert på den dataen, så skal det ifølge datatilsynet ikke være lov. For da har du brukt data til noe den ikke skal brukes til. Men det får dere sjekke, for det kan hende det er andre regelverk rundt dette. For det eksempelet her er bare fra [etaten] som sagt.

I: Ja, for det prosjektet her er jo enda i en tidlig fase, og det kan hende det er ting som må endres.

P04: Ja, men det er viktig. For det her er jo på en måte ting som må bygges på, og det er viktig å ha det klart.

I: Ja, absolutt. Så her kommer det siste spørsmålet. Hvis du hadde vært en foreleser som kunne brukt quizer i undervisningen, hvor ofte tror du at du at du ville ha brukt en slik applikasjon som SmartU?

P04: Da tenker jeg at det måtte ha vært koblet til kursdesign, hvis det er faser i faget som studenter må gå gjennom. Sånn som i metode-faget har vi en teori-modul i begynnelsen. Og da er det naturlig å ha det på slutten av den. Og så har vi kanskje sånne milepæler da. Jeg tror det hadde vært stressende å ha mange slike quizer, jeg tror det må være koblet til et eller annet. Eller kanskje studenter kunne ha laget det for seg selv. Så kunne de ha blitt trent i disse målene, så blir det er mer sånn selvetjening. Som faglærer tenker jeg at det må være tydelig i forhold til planen i faget, hvorfor disse her kommer.

I: Ja, vi har fått veldig mange nyttige innspill, og setter stor pris på at du ville stille til intervju.

Interview participant 06

I: Så først og fremst, hva synes du om å bruke dashbordet?

P06: Ja altså, jeg oppfatta det som relevant for den aktiviteten som jeg driver med, og jeg er jo forsåvidt ute etter måter å automatisere deler av øvingsundervisninga på og kanskje gjøre det mer finmaska med tilbakemelding til studentene også, fordi at – nå så jeg jo ikke på den sida av saken her, men jeg regner med en del av tilbakemeldinga her går det an å presentere for studentene på noe vis også håper jeg da, for at det er jo noe som studentene ofte ønsker seg da på øving. Så jeg ser det som relevant for det som jeg ville ha brukt i undervisninga mi, det gjør jeg.

I: Den studentsida den viser lignende statistikk bare retta mot enkeltstudenter så det er fint at det høres interessant ut. Kan du nevne noen funksjonaliteter som du synes var ekstra nyttig i systemet du prøvde nå?

P06: Det som var mest ulikt det som jeg har sett før det var jo den der statistiksida som presenterte letteste spørsmål og vanskeligste – eller mest feilsvarte og mest feilsvarte alternativ – det er jo noe som jeg har sittet og fikla med manuelt i regneark før, så det er en statistikk som jeg er interessert i da. Og det var jo kanskje på det punktet det som var nytt for meg her i forhold til sånne ting før det var akkurat det at ja akkurat den detaljen der da. For ellers så minner det jo om en del annet, ja jeg nevnte jo Blackboard Collaborate og sånn her i stad, men akkurat det her med å identifisere sånne flaskehalsar det var en ting som jeg før har drevet med manuelt i regneark så det likte jeg.

I: Var det noe funksjonalitet som du synes det motsatte om? Litt mer forvirrende eller vanskelig å forstå?

P06: Det var jo noen sånne knapper og menyvalg som var litt unintuitive at det ikke gikk tydelig fram at her må du dobbelklikke det gikk litt lite tydelig fram at her må du trykke om du skal legge til noe nytt og sånn, men det handler jo egentlig bare om detaljer i grensesnittet sånne som sikkert kunne bli gjort noe med i løpet av ganske kort tid.

I: Jeg så at du trykte ikke så mye på grafene og visualiseringene, men det var sikkert fordi du synes den statistikken var mer interessant?

P06: Ja, også fulgte jeg jo stort sett ordre for å si det sånn da. Jeg fulgte først og fremt ordrene. Jeg er litt sånn induktiv av natur, som liker å finne ut av ting ved å prøve det ut, men det var vel liksom ikke konteksten for det nå da, så da fulgte jeg ordre tross alt

I: Så du rakk ikke å få noe formening om noe visualisering likte ekstra godt eller ikke likte?

P06: Det var...det å bruke sånn bakgrunnsfarge det hender at jeg bruker selv når jeg lager regneark. Og der sørger jeg for å kunne stille på disse terskelverdiene, det går det kanskje an å gjøre her også, å stille inn hvor grensen mellom grønt og gult og rødt går hen.

I: Nei det mener jeg kanskje vi ikke har.

P06: For det kunne jeg kanskje ønsket meg å hatt da.

I: Tror du at du kunne brukt informasjonen på dashbordet til noe nyttig? Til å gjøre endringer i læreplanen eller...

P06: Ja på kort sikt. Det aller mest nærliggende det er jo...det er jo å kunne identifisere ting til, altså temadetaljer til repetisjon og til ekstra øvingsoppgaver og i litt større tidsperspektiv framdriftsplan for semesteret og enda større tidsperspektiv, altså selve emneplanen da. Så det er flere muligheter her, som det kunne vært en nyttig tilbakemelding.

I: Enn det å nå ut til studenter som det ser ut til er i fare for å stryke eller gjøre det dårlig: tror du at du hadde kommet brukt det til noe sånt?

P06: Det har jo veldig mye med instituttets politikk på sånne ting og hva en skal prioritere og sånn. Selv så har jeg litt sånn blanda erfaring med hvordan studentene reagerer på å bli tatt kontakt med på den måten da. Det er jo noe som alle vet at noen studenter liker å ha **banker bortover på bordet** en fast og streng fremdriftsplan, mens noen studenter liker å ha det fritt og kunne dra til Åre på en torsdag ikke sant. Så det er fra individ til individ.

I: Hva syns du om designet til dashbordet? Var det overveldende eller var det litt lite?

P06: Nei jeg hadde ingen problemer med å takle det. Jeg har jo nevnt flere ganger at det var noen sånne knapper og menyvalg som var litt misvisende eller litt lite utfyllende. Bortsett fra det så syns jeg ikke at det var overfylt det syns jeg ikke.

I: Jeg vet ikke om du la merke til det, men jeg hørte at du nevnte det på slutten at du så et rødt utropstegn...

P06: Ja! Jeg hadde – det var akkurat på tidelssekundet at jeg så det røde utropstegnet akkurat samtidig som fingern var på tur ned på musa **peker med fingern ned på datamusa**.

I: Så du det røde utropstegnet når du var inne på dashbordet også?

** viser frem varsler inne på dashbordet **

P06: Altså jeg tror ikke at jeg...jeg så liksom mer på kodinga ikke sant **gestikulerer til et spørsmål med programmeringsinnhold som er på skjermen**. Der ja, jeg var jo inne å så på den der, men jeg klikka ikke på utropstegnet i seg selv. **Holder musepekeren over varslingsikonet**. Men der kommer det fram en melding ja.

I: Men la du merke til at det var der første gangen når du var der?

P06: Nei jeg tror ikke jeg reflekterte over det. Det var sånn tilleggsinformasjon som jeg kobler ut første gangen som jeg kanskje ville sett på andre eller tredje eller fjerde gangen. For her står det noe om at det er irrelevant...nei det står at valg-raten er over terskelen. Men hva betyr egentlig terskelen? Har det noe med... jeg snakka litt om grenseverdier mellom farger i sted, for jeg regner med at det er noe som ikke er større enn 100% for eksempel, noe som åpenbart er feil, det er ikke det det refererer til men det er heller det at det er dårligere enn det du hadde regna som dårlig eller et eller annet sånt noe. Så her er det kanskje noen sånn flaskehals som vi burde se nærmere på altså.

I: Nå som du vet dette da, tror du at det hadde hjulpet deg med å se hva du burde være oppmerksom på, at utropstegnene kanskje kan hjelpe deg med å...ja?

P06: Jeg tror det er mer spørsmål om tid jeg. Å bruke mer enn 7-8-9-10 minutt, jeg tror det hadde hjulpet ganske mye bare i seg selv jeg. Jeg ville nok ha klikka på det, eller sett på det utropstegnet om ikke så lenge for å se, men jeg filtrerte det litt bort nå i sted for da var det mest om å gjøre å navigere seg gjennom den løypa som du la opp.

I: De studentlistene du ser, for eksempel til grafen, de fargene som er satt der med terskel med stryk med 40 poeng, **viser elementer med nevnte farger på skjermen** nå er de sortert på siste forsøk på quizen de har tatt, så den scoren de har fått på den siste quizen. Men vi –

P06: Vil det si...skal den grafen der leses som at de har hatt seks ulike etter hverandre da?

I: Det varierer hva studentene har hatt. Den blå streken er gjennomsnittet, og så skal du kunne trykke på en så kan man se for hver enkelt student. Men det vi er usikre på er hva som er interessant for en lærer å se: om man ville hatt en sortert liste på gjennomsnittlig score eller en utregning fra de siste forsøkene, eller som vi har gjort det nå at vi bare viser siste forsøk?

P06: Det som er litt interessant, det er jmføre aktiviteten til slutten av semesteret med gjennomsnittet for eksempel da. At det er på en måte...det er kanskje mindre farlig om det er noen som har vært slappe i starten og som har kommet seg på slutten enn omvendt da. For det siste er et mye større faresignal enn det første. Det er å dra ut en form for deskriptiv statistikk, et tall som sier noe om kurven går oppover på slutten eller nedover på slutten.

I: Så et avsluttende spørsmål. Tror du at du ville brukt en sånn her applikasjon om du hadde hatt muligheten til det?

P06: Ja, alt må jo sees i en kontekst, men jeg ser det absolutt som ikke uaktuelt, for det er en del funksjonalitet her da som jeg ikke har sett i andre systemer som jeg egentlig har ønsket meg da. Så bare det i seg selv er et argument. Også hadde du dette her med at du har spørsmål, en bank med spørsmål eller ei fil med spørsmål, det går jo an i det systemet som vi har nå også, Blackboard, men nokså kronglete og veldig vanskelig å eksportere data fra og eierskapet til spørsmålene kan bli litt tvilsom også, og dersom NTNU bestemmer seg for å bytte ut systemene etter 4 år hva da? Jeg så at det var .csv og det tyder jo på at det er et veldig generelt og grunnleggende tekstformat, det lover jo godt for å kunne overføre dette til et annet system kanskje. Jeg liker at ting er mest mulig åpne og fleksible på den måten.

Interview participant 07

I: How did you feel about using the SmartU dashboard?

P07: Very comfortable. I would say that it is really easy to go through and understand what is the level (of the class/students) and several other things. I would say it is really complete. In general I would say it is something that is easy and comfortable to use.

I: Could you mention a few features you found extra useful and explain why?

P07: Extra useful... in a bad way or a good way?

I: In a good way!

P07: Ok. I would say that it's really cool that I can edit and see what is good and not good. And also the progress of the students and that you can compare them. I mean I love everything! I like the student overview and also the question distribution so that I can see what I can improve or what I can edit or add later. All these features I find really useful, but let's say that the extra usefulness is more about those parameters (pointing at the statistics/activity insights). Because they are very immediate and give me the potential to improve in a really fast way and I love it.

I: Were any of the features confusing or difficult to understand?

P07: I would say that the mastery grid and the filters could be improved a bit. I know that with 50 students it is a bit difficult to explore through all 50 students, but it is useful. But it could be improved in a way where I can also make a comparison between different students. I don't know if it is possible?

I: No, it's not possible.

P07: Ah, ok.

I: What was your most preferred visualization in the dashboard, if you could mention one, and why?

P07: Hmm... That's a good question. I do not have a preferable one. I would say that this is very useful for a teacher **pointing at the learning progress graph**. So this is something I would look at a lot. And I also like the overview. This one is also cool **pointing at the mastery grid**, but it is much more complicated and not as easy to understand at first glance.

I: So you would say that the mastery grid was the least preferred visualization?

P07: Yes. Not only because it is not readable, but just because it needs some more time to read.

I: Are there any other features you wish were present on the dashboard?

P07: The one thing that I told you.. the save things, but it is not a feature, just something to add. And the other thing I talked about, comparing students would be cool to have as a

teacher. If a teacher wants to understand how and why. If it is a combination of multiple difficulties. But for the rest it is very complete.

I: How do you think you could utilize the information presented on the dashboard?

P07: Utilize for setting up the next questions related to the curriculum, but in a more customized way. This dashboard will for sure help me understand the level of the students and how to design the next activity. I would say this is very powerful and something that is missed in schools and universities, to help teachers understand how or why a task is difficult or not. And to then be able to modify it.

I: How did you feel about the notifications/highlighting? Did you find them overwhelming? Did they help you focus on specific elements in the dashboard?

P07: The notification/highlighting helped me to understand when an action by me was required and think this is really helpful, they weren't disruptive.

I: What did you think of the dashboard's design?

P07: I love it. I think you did very good work in terms of the visualizations. I can easily reach what I want. For sure it can be improved a bit. It could also have some customizable things, but that will be another level. I am happy with the design and it gives the perception that it is very easy to use.

I: I saw you did play around with the filters a bit. How do you prefer to see student scores displayed in a system like this where their scores are updated with each new attempt? Is the last score the only interesting metric or do you prefer an average score of all attempts or a few of the last attempts?

P07: Hmm.. Yeah, maybe this is even more difficult, but having a range of time which I can select. I don't know. The first month related to the last would be more useful for me to better understand the average of the score. It is even more complicated, but I think if you design in a good way it is very useful to have the possibility to select a range of time. It is useful to know the last attempt to know the current progress, and this could be the first information you gave me. And it is perfectly OK. But if you want to go in more deep and investigate further it would be interesting to be able to select a range of time.

I: Thank you, we didn't even think about that solution! That sounds like a really smart feature. And to the last question: If you had access to the SmartU application, how often do you think you would use it?

P07: Always! I would say as a teacher... Once I start creating all the activities that are needed, I would love to continue on this path. Once you start you can collect information and monitor all the activities and students in a really good way. Maybe I'm not the average teacher, but if I think about a teacher from school and not from university, maybe an initial point could be a bit difficult for them, since they are using some panels in terms of design in the activity and so on. But I think once they understand the potential around it they will use it. If it were up to me I would use it every day.

Interview participant 08

I: Hva synes du om å bruke SmartU-dashbordet?

P08: Nei det var bra, nice, det var enkelt å bruke, intuitivt. Hadde, ja, stort sett den informasjonen jeg ser for meg er nødvendig.

I: Kan du nevne en eller flere funksjonaliteter som du synes var ekstra nyttig og hvorfor?

P08: Ja...tenker kanskje den oversikten hvor du kunne se den progresjonen til hver student over tid og over forsøk. Den virka nyttig. Blant annet, for det er lett å se om de klarer å...om det står helt stille da eller om man faktisk forbedrer seg.

I: Var det noe funksjonalitet som var forvirrende eller vanskelig å forstå?

P08: Nei jeg tenker ikke det. Det var den ene tingen med liksom når man hadde et bilde til et spørsmål så ville jeg kanskje flytta det opplaste-bilde-boksen bare rett under bildet.

I: Sånn at det er en sammenheng mellom bildet og boksen?

P08: Ja, det er kanskje det eneste. Jeg vet ikke om det er bedre heller da, men jeg bare så for meg at kanskje det er bedre.

I: Nå svare du jo egentlig på at du synes den progresjons-greia var best og det er egentlig det neste spørsmålet: hvilken visualisering som du likte best, så vi hopper over den. Men var det noen av visualiseringene du ikke likte eller som du kanskje synes var litt dårligere enn de andre?

P08: Nei, jeg tror ikke det alt var veldig tydelig.

I: Så du tror du kunne ha brukt alle og tydet noe informasjon ut av det som du kunne ha brukt til noe?

P08: Ja. Ja helt klart. Helt klart.

I: Var det noe funksjonalitet som du kanskje skulle ønsket var tilstede på dashbordet? Noe du savnet?

P08: Nei nå vet jeg jo ikke helt hvordan det her funker da, hvordan man skal bruke det og sånn. Men kanskje sånn hvis man hadde veldig mange klasser og man hadde en student som strevde eller what ever i hver klasse og så hadde man...hvis jeg liksom hadde iverksatt noe da, så kunne kanskje vært nice å bare ha et sted hvor jeg kunne notert det liksom.

I: En egen notatside?

P08: Ja, at jeg har tatt tak i. Til den enkelte studenten. Sånn hvis jeg bare skal finne på et eller annet, men det kommer veldig an på hvordan det brukes selvfølgelig. Kan hende det er naturlig å gjøre det et annet sted liksom.

I: Tror du at du kunne ha brukt informasjonen som ble presentert på dashbordet til noe nyttig, i den arbeidssettingen som stud.ass.?

P08: Ooh, som stud.ass.? Ja kanskje ikke akkurat med de arbeidsoppgavene vi hadde nå da. Men det kan hende det er litt sånn forskjellig fra fag til fag. Så jeg kan se for meg at det er nyttig for noen, men ikke nødvendigvis sånn som vi gjorde det. Hvis for eksempel et ansvarsområde for stud.assene. er bare å følge med på folk og si ifra om det er noen som strever og sånn, da hadde det definitivt vært nyttig. Og hvis det var sånn at vi som er stud.asser. hadde ansvar for å oppsøke folk som strevde med å få hjelp, eller med å hjelpe dem, så hadde det vært nyttig. Ja, men det liksom ikke helt, sånn som det funker nå da, men det kunne kanskje være en bedre måte å organisere det på eller noe sånt no.

I: For vanligvis ville det vært slik at de oppsøker dere? Og de har kanskje andre typer oppgaver de løser enn flervalgsoppgaver?

P08: Ja, ja riktig. Og ja det er forsåvidt sant. Man kunne lagt opp ting annerledes da, til å definert arbeidsoppgavene til stud.assene. litt annerledes.

I: Litt i samme gate: hvordan tror du at dashbordet kunne hjulpet deg med å nå ut til studenter som systemet da mener står i fare for å stryke?

P08: Ja jeg tenker bare det å få identifisert dem, da er jo veldig mye gjort.

I: Og hvordan tror du at du ville ha tilnærma deg den situasjonen?

P08: Ville laga en general announcement tenker jeg også publisert at [Ola Nordmann] strever. Neida hehe. Jeg tror kanskje...nei jeg er ikke helt sikker. Men det er liksom på en måte hvis du vet hvem det er da, så kan du jo kontakte på et litt mer diskret hvis da.

I: Nå har kanskje ikke stud.asser. så mye makt til å gjøre endringer i læreplanen, men hvis du hadde hatt det, hvordan tror du at du ville brukt dashbordet til å hjelpe deg med å gjøre passende endringer?

P08: Jeg vet ikke, du får jo innsikt i hva det er folk strever med og sånn da. Det kan man jo gjøre på andre måter også, men det er jo en veldig sånn effektiv måte å se hva kanskje man skal bli bedre på. Ja hvis for eksempel noe man antar er veldig lett da, som de – som det er mange som svarer feil på, så er jo det innsikt om at det kanskje, at man burde lære bort på en annen måte. Og vice versa, hvis det er noe som skal være veldig vanskelig, men alle får det til, så kanskje det ikke er så vanskelig.

I: Hva synes du om designet til dashbordet?

P08: Nei jeg synes det var fint. Lett å bruke, oversiktlig og masse informasjon på liten plass uten at det føltes overvelming.

I: Ja det er bra, for det er oppfølgingsspørsmålet, om det enten for simplistisk eller for avansert. Du følte ikke på noe tidspunkt at "her var det for mye" eller "her skulle jeg gjerne hatt mer"?

P08: Nei, veldig bra.

I: Jeg vet ikke om du la merke til det, men vi hadde noen sånne notifications som var når du starta, både på aktiviteten der hvor du velger aktivitet og inne på de "most failed question" og på noen studenter i lista. Hva synes om de?

P08: Nå er det sånne notifikasjoner som forsvinner etter du liksom klikka på den?

I: Ja.

P08: Hmm, ja det vet jeg ikke, det kommer litt an på hvordan de fungerer. Jeg vet ikke om jeg skjønnte hvordan de fungerer enda. Men hvis det er sånn at hvis noe har endra seg så tydeliggjør den det for eksempel så er jo det kanskje fint.

I: Følte du at de hjalp deg med å rette blikket mot...?

P08: Ja helt klart.

I: Tenker du at du kanskje ikke ville sett på en student dersom de ikke hadde en sånn varseltrekant for eksempel?

P08: Ja det blir iallefall veldig lett å se på de som har det, det er det ingen tvil om.

I: Så ville du da foretrukket med eller uten disse notifikasjonene?

P08: Nei jeg vet ikke, det kommer litt an på hvordan de genereres. Så hvis de liksom, det står en notifikasjon når noe endres så tenker jeg det hadde vært bra. Men hvis denne notifikasjonen er der selv om jeg har sjekka det fem ganger på rad og den bare alltid gir meg notifikasjon på de som gjør det dårlig så tror jeg at jeg ville space de ut litt etterhvert. Så det kommer litt an på hvor smarte de er.

I: Du kikket litt på filtrene til studentlistene, og de er i utgangspunktet basert på siste score. **viser filtrene på skjermen**. Så her ser vi studentens score fra siste attempet som de tok. Hvordan ville du foretrukket å se studentens score vist fram i dette systemet? Har du noen tanker rundt det å se score fra alle siste, eller average, eller tre siste, siste to uker...har du noen tanker rundt det?

P08: Ja, nei jeg tenker jo at det var nice sånn som det er nå da. Jeg kan jo se ut utviklinga på noen andre. Ja jeg vet ikke, det er jo... jeg tenker det er greit sånn som det er nå kanskje. Kanskje liksom hvis jeg hadde, hvis jeg kan klikke meg inn på den boksen **klikker inn på en student i mastery grid-lista**. Ja ikke sant for da er det ikke noe pes, jeg tenker det er nice sånn som det er nå.

I: Ja du kan liksom trykke deg inn på denne her **trykker inn på en student som har scoret 0 poeng i en kategori og velger deres siste forsøk** Da ser du at han bare har fått ett spørsmål om binær. Kanskje han har vært god på det tidligere. Endrer det noe på hvordan du tenker om det?

P08: Nei det er veldig oversiktig sånn som det er nå.

I: Okei, aller siste spørsmål: hvor ofte ville du brukt en slik applikasjon?

P08: Nei det er sikkert noe man ville sjekka hver gang man skulle ha en sånn hjelpe session da, så samme som hvor ofte det er da.

I: Så om det var en øvingstime i uka ville du brukt det en gang i uka?

P08: Ja.

Interview participant 09

I: Ok, så første spørsmål er hva synes du om å bruke SmartU dashbordet?

P09: Veldig mye mer forståelig enn for eksempel Blackboard navigasjon, det må jeg si. Det er det eneste referansepunktet jeg har her da er jo Blackboard. Og det er jo veldig ur-brukeropplevelse synes jeg, i forhold til det her som er moderne cards-basert. Ganske lett å – **klikker seg inn på en aktivitet** – med en gang her så får jeg ganske mye oversikt. I stedet for å drive å grave meg gjennom meny etter meny. Det synes jeg er positivt.

I: Kan du nevne en eller flere funksjonaliteter som du synes kanskje var ekstra nyttig og hvorfor?

P09: Ja, jeg liker det med filtrere og bare kunne sortere på scores her. Jeg tenker jo at jeg da i min stud.ass. mind tenker jeg jo veldig sånn katastrofehåndtering, ikke nødvendigvis å se at det går så veldig bra, det er spesielt fokus på der det ikke går bra tenker jeg. Så det er bare det med å se, spesielt opp i mot en person og så bare se hvordan de ligger an per øving eller test eller hva jeg skal kalle det. Og så det her med commonly failed, greit nok, er det ordlegginga eller er det bare at det er et vanskelig spørsmål, såne der ting man kanskje må ta stilling til som stud.ass. Men jeg kan notifisere læreren da om at "se her det her funker eller er dårlig".

I: Er det noe funksjonalitet som var forvirrende eller vanskelig å forstå?

P09: Det første jeg tenkte på er opplastinga da. Det er jo sånn at når jeg har gjort det en gang så forstår jeg det. Det er minimal opplæring. Hvis jeg hadde sett liksom en film eller en liten snutt så hadde jeg nok skjønnet det. Men new activity her, så kan jeg skrive inn noe, save and exit **viser til knappene "save and exit" og "next" på new activity-skjemaet**. Man blir jo ikke umiddelbart tatt til – for det er jo "save and exit" – så kommer man hit også er det den data prompten her **peker til hjelpetekst**. Det var kanskje bare ikke helt tydelig for meg umiddelbart før opplæring at create a new activity hadde sammenheng med filopplasting bare ut i fra dette her. Også kanskje bare plasseringa, notification jeg trykka på den, den var iøynefallende, så kom jeg på dashbordet og det var notifications så jeg på de her **peker til activity insights** som sikkert var nested, sånn at den på toppnivå fikk notifications fordi den her hadde notifications. Men jeg klarte ikke å koble det umiddelbart, for det var så mye farger og fint på toppen. Men selvfølgelig hadde jeg vært kjent med systemet i mer enn tjue sekunder så hadde det sikkert vært veldig lett.

I: Spesifikt om visualiseringene, var det noen der som kanskje du synes var bedre enn andre?

P09: Jeg synes alt bare var fint jeg.

I: Så det var ikke noen som enten var verst eller best?

P09: Nei. Altså jeg liker de her mastery grid. Ja, jeg tror kanskje det er den mest innsiktsfulle biten synes jeg, for krisehåndtering skulle jeg til å si. Så det var fint at den var fargekodet og lett å se.

I: Du hadde jo en del interessant idéer underveis mens du holdte på, så spørsmålet er hvilken funksjonalitet som du skulle ønske var tilstede på dashbordet?

P09: Det er jo noe med – hva heter det, scope creep? Jeg vil ha det, jeg vil ha det, jeg vil ha det. Men, kanskje bare noe sånn e-post med en gang jeg trykker på Tuva her, så får jeg opp hennes epostadresse her, men du vet det finnes jo sånne linker osv. som du trykker også bare åpner den opp Outlooken din. Det er kanskje en fiffig, nyttig ting hvis jeg vet at det finnes da. Så det kan jo kanskje være indikert med en e-post her hvis ikke det clutterer opp viewen. Så det er jo en sånn funksjon det hadde vært fint å ha. Direktelinje til lærer er jo litt det samme, men hvor det ville passet best inn, altså det er jo fint å kunne gi referanse direkte til hva, altså jeg kan skrive i e-posten til læreren “obs obs most commonly failed question 7 må du se på”, men det er kanskje å på en måte linke til direkte her, er det mulig? Å få direkte link?

I: Det er vel foreløpig ikke mulig. Da måtte du i så fall bare skrive hvilket spørsmål det var snakk om i en e-post.

P09: Så det er kanskje det da å kunne linke til spesifikke objekter å se på de.

I: Så nevnte du også når du var inne å så på de tidligere forsøkene til studentene, du ville se spesifikt hvordan spørsmål de hadde svart på også?

P09: **viser til et forsøk på en student** Ja for jeg ser her, det er noe med algoritmer som har skjært seg, men akkurat hvilke spørsmål den fikk kontra noen andre det lurer jeg på. Hvis det var noen id eller noe sånt.

I: Tror du at du kunne brukt den informasjonen som ble presentert på dashbordet til noe nyttig?

P09: Ja. Hvis min jobb hadde expressly det å – nå ble det mye engelske ord – hvis det hadde vært uttrykkelig at min stilling er å passe på å stryke, på det nivået, så hadde det her vært veldig nyttig tror jeg. Spesielt det med sånn tids... jeg kan se progress by attempts og så videre. Det er litt sånn jeg kan ta litt grep der da. Se de som prøver mye men ikke får til, kontra de som ikke har prøvd i det hele tatt, og distribuere min innsats etter den som prøver mye for eksempel hvis du må gjør noe sånn triagering. Så jeg tror du kunne reddet mange strykkarakterer med et sånt system.

I: Du har egentlig besvart det neste spørsmålet, om du tror det kunne hjulpet deg med å hjelpe studenter som står i fare for å stryke. Og du ville foretrukket e-post eller å snakke med de eller?

P09: Ja. Ja, det er jo den beste måten tror jeg. Det er kanskje litt forceful med å kontakte direkte. Og det er jo noe med, jeg kan jo gå inn å endre spørsmål, kanskje ikke jeg som stud.ass. skal få til det, men det går jo an å legge opp til at noen skal lykkes ved å skru ned vanskelighetsgraden. Det er jo kanskje ikke riktig strategi, men det kan jo hende. Så most commonly failed question da, jeg vet ikke hvorfor det er most commonly failed. Så det går kanskje ikke an å få noen innsikt via kanskje bare statistikken. Det kunne, hvis studenten hadde fått anledning til å kommentert, det kunne vært et artig datapunkt å hatt i tillegg, “dette spørsmålet synes jeg var veldig forvirrende”. Litt tilleggsinformasjon.

I: Som stud.ass. har du kanskje ikke anledning til å gjøre endringer i læreplanen, men tror du at du kunne sett over tid noen endringer som måtte gjøres ut i fra det du ser her?

P09: Så jeg, på bakgrunn av denne her dataen kunne gjort inngripen? Ja. Det har jo sagt noe om, altså du må jo korrelere det med annen data også sikkert. Altså oppmøteprosent i timene og sånn og hvorfor, hvor problemet ligger. Men det er i alle fall et nyttig datapunkt da i den blandinga tror jeg, å kunne se hva som, altså hvis jeg kunne se akkurat hva spesifikt. Jeg vet nå at algoritmer er en utfordring for mange **viser til mastery grid** ser det ut som. Ja det er jo også en ting her, kan jeg se kun algoritmer, altså kan jeg sortere etter...?

I: Det går nok ikke an akkurat der.

P09: For det er jo sånn, kan jeg vite antallet som stryker per kolonne her så hadde det også vært interessant. For da kunne man lagt opp i læreplanen og gjort endringer der.

I: Hva synes du designet til dashbordet? Var det overveldende, akkurat nok, for lite?

P09: Jeg synes det er godt for hva det er, altså at du har en stor graf her **viser til progresjonsgrafen**, det er bra at den har fått nok plass. Og det er fint at alt er – jeg vet ikke om jeg skal begynne å **endrer vindustørrelsen** – joda det er jo, hva heter det flex? Jeg tror det er nok plass og at det er spacea greit. Det er det med, toppen her synes jeg er veldig iøynefallende, men det er også mitt førsteinntrykk da, men så så jeg at det var notifications på de her. Så blikket ble jo fanget av de lyse fargene syns jeg på toppen. Og det kunne kanskje ha stykka det opp litt i høyden, eller kategorisert det da, alt er jo sammensatt data forsåvidt. En og en ting hadde gått an kanskje, så en bolk med analytics, kanskje de her **peker på activity insights** overordna på toppen eller noe sånt. Eller overview selvfølgelig. Så en og en, A-B testa med det.

I: Du nevnte det litt nå med at det var notifications på de mindre analyticsene. Hva synes du om notifikasjonene som du så i systemet? Både den på forsiden og inne på aktiviteten?

P09: Så jeg så jo at det var et utropstegn, og tenkte bare at "ok det er noe som trenger min oppmerksomhet". Og jeg antar at de feltene som hadde notifications her **peker på activity insights** har blitt oppdatert siden sist gang jeg så det. Jeg åpna de også forsvant de, men hvorfor var det notifications in the first place tenker jeg da, som krever min oppmerksomhet? Er det det at – det er sånn call to action spørsmål da, for jeg tenkte at her må jeg ta stilling til at mange failer question 7, og den informasjonen er ny antakeligvis siden det var notification her. Så det var jo hvis jeg forstår arbeidsprosessen ellers så hadde det jo vært veldig nyttig.

I: Så du foretrekker å ha det kontra å ikke ha det, har du noen preferanser der?

P09: Ja, det er bedre å ha dem ja. Altså hvis de her systemene hjelper meg med å sortere ut. Og spesielt det å skjønne at det er nytt siden sist liksom. Men spørsmålet er, er det nytt eller er det noe å ta tak i? Det kunne jo vært forskjellige typer notification-ikoner, om det bare er nytt eller om det er noe ekstra.

I: Du var inne på de filtrene og prøvde litt forskjellig sånn sortering, som standard så er det sortert etter scoren deres på siste forsøk de hadde på quizen. Har du noen tanker om

hvilken score som du synes gir mest mening å vise fram? Er det en average av alle, eller siste, eller tre-fire siste eller...?

P09: Jeg tenker at jeg er mer kritisk til sånn eldre – altså jeg vet ikke helt hvordan læringsoppsettet her er da – for den last attempt, er det last attempt på Python, General Programming, hvilken last attempt er det snakk om?

I: Det er bare en quiz, hvis du for eksempel går inn på Tuva **trykker inn på Tuva og velger siste attempt**, så kan du se attemptsene og hvilke typer questions de svarte på. Så scoren er ut i fra det som står der.

P09: Mhm, så da må jeg kryssreferere det. Så jeg trykker på Tuva nå, så trykker jeg på Algoritmer **trykker på score cellen på Tuva sin rad, kommer til samme dialog som tidligere**, å ja det er samme for hver.

I: Ja, det åpner samme dialogen.

P09: Ja, for det kunne jeg kanskje tenkt meg å...nei jeg vet jo ikke egentlig hvordan prøven ser ut da, om de tester dem i Algoritmer, Binary og General Programming og Python i en prøve?

I: På grunn av at den [quizen] er adaptiv så er det ikke garantert at du får spørsmål fra alle.

P09: Å ja, men du vil alltid prøve å mikse dem sammen? Men spørsmålet var om det var...?

I: Hvordan data du vil se som deres score, om det er siste attempt eller average av alt eller?

P09: Om de har klart det så er jeg på en måte ikke interessert i de lengre. Kunne sikkert fjerne alle de som har klart det, men de som ikke har klart det, det er spørsmålet da om de har prøvd mye og ikke klart det, eller er det bare at de prøvde en gang og bare ga opp liksom. Så da tror jeg kanskje average er mest interessant for min del, men det er litt nice å kunne vite hvor mange ganger de har prøvd bare med en glance og da.

I: Siste spørsmål, i den arbeidssettingen som stud.ass. hvor ofte tror du at ville brukt en sånn her type applikasjon?

P09: Hvis det her er dashbordet på den arbeids – da går jeg opp her ser om det er noen som trenger hjelp som ikke kontakter meg i tilfelle jeg ikke sitter i rommet. Så tenker jeg “ok han her kan jeg plage” så treffer jeg han og spør om det går bra. Jeg vet ikke egentlig om det er prosedyre da, men det er i alle fall en effektiv måte å få han til å ikke stryke på. Så det er jo å spørre: “trenger du hjelp? Jeg er tilgjengelig, så skal vi komme oss gjennom det.” Så da tror jeg du ville fått veldig få stryk. For jeg antar jo at folk gir jo opp da hvis de ikke får det til etter mange forsøk. Så du berger vel litt der. Jeg tror jeg ville brukt mye tid her.

Interview participant 11

I: Hva synes du om å bruke SmartU dashboardet?

P11: Jeg synes det var veldig hjelpelig. Det ga veldig mye innsikt. Veldig mye. Så jeg synes... det tok meg litt tid å forstå hva jeg skulle se på og hvordan jeg skulle tolke det. Men etter hvert så fikk jeg.. det var veldig lett å bruke det. Altså hvordan det går med elevene.

I: Du sier at det var litt mye og at det tok litt tid å sette seg inn i det. Synes du det var for mye?

P11: Jeg anser meg selv som litt datakyndig, og jeg brukte veldig lang tid på å forstå hva fargekodingen var. Det står jo at det er high risk og low risk. Et annet sted sto det high risk at failing, men det gjorde det ikke her *peker på overview*. Så jeg skjønnte ikke helt hva det indikerte. Ellers så synes jeg det var veldig nyttig med den grafen for da ser du over tid. Mastery grid er også veldig greit, men her ville jeg gjerne sett gjennomsnitt også. Det er mulig det allerede finnes, men at jeg ikke husker hvor det var. Men det som kanskje var best var den statistikken, highlightsene, for da kan jeg ta tak i det med en gang. Mens det andre må jeg bruke litt mer tid på for å skjønne hva jeg kan gjøre med enkelt-elever, tema eller spørsmål.

I: Bare for å få det klart, når du snakker om gjennomsnitt, mener du å se gjennomsnitt for hver kategori for alle elevene samlet?

P11: Ja, eller, hvis jeg vet at algoritmer er vanskelig for en gruppe studenter, eller det å kunne skille på... Det er jo en ting det å se det mønsteret på en gruppe i en klasse, de som gjør det bra eller dårlig. Men så er det også nyttig å vite om... her måtte jeg for eksempel gjøre sånn for å få et bilde over hele klassen. Og da ser man at her er det en gruppe som gjør det veldig, veldig dårlig. Hvordan kunne jeg ha fått tak i... Et gjennomsnitt ville ikke hjulpet meg i det hele tatt. Men jeg vet ikke helt hvordan jeg skulle ha fått tak i...

I: Kunne du kanskje ønske å kunne sortere per kategori, slik at du hadde fått alle de dårligste studentene i en bestemt kategori på toppen?

P11: Ja, eller det var det jeg lette veldig lenge etter. Per tema kunne jeg ha fått sånn at jeg kan ta tak i et tema. Det er det jeg kan gjøre hvis jeg skal undervise. Jeg kan ta tak i et tema, og det må være fordi enten en gruppe av klassen eller hele klassen synes det er vanskelig eller at det er enkeltpersoner som trenger ekstra tilpasning eller ekstra ressurser, andre måter å forstå innholdet på.

I: Ja, dette er veldig nyttig informasjon for oss! Kan du nevne en eller flere funksjonaliteter som du syntes var ekstra nyttig og hvorfor? Det så ut som om du likte statistikkene?

P11: Ja, de var veldig lavt hengende frukter. Jeg så de med en gang, så derfor begynte jeg ikke å trykke på de med en gang, men utforsket heller de andre elementene først. Statistikken er iallfall veldig lett og det at jeg til og med kan edite direkte i dashboardet...! Det er ganske smooth. At jeg ikke trenger å gå til

question editoren. Den question editoren synes jeg var verre, for det å skjønne at man må trykke nok ganger for å sortere var mindre logisk. Statistikken var iallfall veldig nyttig. Jeg slipper å forlate dashboardet for å edite spørsmålene da.

I: Det er bra! Var noen av funksjonalitetene forvirrende eller vanskelige å forstå?

P11: Det må bli fargene da.

I: Det kan jeg forresten klargjøre med en gang. Jeg ville ikke blande meg inn under brukertesten med mindre du spurte om hjelp. De fargene er basert på NTNUs prosentvurderingsmetode, der en poengsum på 40 eller mindre regnes som stryk og dermed rød, mens dårligere enn C, mindre enn 65 poeng, er gul, mens C eller bedre er grønn. Det er ikke sikkert det er den beste måten, og vi skulle iallfall ha vist hva fargekoden betyr tydeligere i dashboardet.

P11: Hvis jeg hadde vært på NTNU hadde jeg kanskje antatt det. Jeg synes ikke det var intuitivt hva lav, medium og høy risiko betyr. Ja, og en ting til! Jeg synes også det var vanskelig å finne fram filtrene og skjønne forskjellen på last attempt og average. For da fikk jeg sett det hele bildet. Kanskje last attempt er vel så viktig.

I: Vi har faktisk et spørsmål og akkurat det der, så da kan vi like gjerne ta det med en gang nå som vi er inne på temaet!

P11: Ja, ok!

I: Vi er jo ikke lærere, så vi er ikke sikker på hva som er nyttig for en lærer å se. å vi var usikker på hvordan score burde sorteres. Ettersom en average score kan gi et mye dårligere resultat enn studentens nåværende nivå pga. tidligere dårlige forsøk, har vi satt last attempt som standard visning. Det finnes jo også andre alternativer som f.eks. sortere på en score som er basert på de siste 3-4 attempts o.l.. Har du noen tanker rundt hvordan du hadde foretrukket å se score bli sortert?

P11: Tja, det er litt forskjellig. Average er nok det jeg hadde brydd meg om når jeg ser på og skal forstå pedagogisk hvorfor det er utfordrende, det temaet eller kategorien vi går gjennom nå. Det at du får 6 forsøk, det er ikke alle som kommer seg til forsøk 2, 3 og 5 hvis du skjønner. Men jeg vet ikke... Jeg blir litt forvirret. Er det sånn at alle må ta alle 6 forsøk?

I: Det varierer, studentene kan ta så mange forsøk de vil. Så en student kan bare gjøre 2 forsøk, mens en annen kan gjøre 20 forsøk.

P11: Hmm... Nei for jeg er jo enig i at jeg bryr meg om foreløpig status, hva kan studenten nå eller ikke, og det er jo last attempt. Men hvis man har forsøk som viser at du må ta testen 6 ganger for å forstå det, så må det være noe galt i måten... hmm, det er noe informasjon der da! Jeg prøver bare å tenke hvis jeg skulle ha hatt en time med videregående og funnet ut av hva noen synes er utfordrende. Da ville jeg sett på... kanskje det blir riktig med last attempt likevel? For da er du motivert til å finne ut av det. Det kan stemme at last attempt faktisk ikke er så dumt.

I: Hehe, det er ikke bare vi som synes det er vanskelig å finne ut av hva score burde sorteres etter.

P11: Nei, det høres riktig ut. Det var ikke akkurat å fryktelig godt svar, men det var det jeg kom fram til!

I: Nei, det er godt nok for oss det! Og så litt tilbake til de statistikkene igjen. Jeg vet ikke om du så når du åpnet systemet for første gang, så vet jeg ikke om du så slike røde dotter?

P11: Jo.

I: Hva tenker du om slike notifikasjoner? Vil det hjelpe deg med å se hva som man burde ta en titt på?

P11: Jeg begynte jo ikke med å trykke på dem, vet ikke hvorfor jeg er sånn. Men ja, det gjorde jo at jeg visste at jeg skulle innom der uansett. Så... det var jo også notification på forsiden, som sier meg at det er noen oppdaterte verdier. Noe sånt er jo veldig greit å få beskjed om. Samtidig burde det jo være... Det kan være smart og behag da, men det burde jo være det første jeg sjekker ut om jeg synes det var bra. Hmm.. nei, jeg vet ikke.

I: Men synes du det var forstyrrende på noe vis da?

P11: Hmm... Nei, jeg ville ikke si at jeg synes det var forstyrrende. Jeg vil ikke si at jeg ikke så på det først pga. det var av protest. Det var rødt med utropstegn hvis jeg husker riktig?

I: Ja.

P11: Så det kan jo også virke som systemkrasj... Men det var samtidig egentlig helt greit.

I: Da må jeg bare grave litt mer jeg! Hvis du hadde hatt valget om å ha notifikasjoner eller hatt muligheten til å slå av i settings, ville du hatt det av eller på?

P11: Hvis jeg hadde hatt flere klasser så... Tror ikke at jeg hadde slått det av. Kanskje etter hvert, men i starten når jeg enda ikke vet ikke hvordan det fungerer så ville det vært greit å få litt hjelp til hvor jeg burde se.

I: Ja, det høres greit ut det. Og så lurer jeg på om du kunne ha brukt noe av informasjonen her til å gjøre noen endringer i undervisning eller læreplan?

P11: Ja, undervisning spesielt. Jeg antar dette er adaptiv quiz man jobber med etter undervisning, så toget har jo på en måte gått med tanke på tilpasning undervisning. Men det man kan gjøre er å adressere temaene man synes er vanskelig ved å legge ut ekstra ressurser på ulike format. Folk snakker om at de er visual learners og forskjellige ting, og det stemmer ikke. Men det å ha forskjellige måter å presentere

det på... Men det krever veldig mange ressurser, Og det burde man bare gjøre hvis det er behov for det. Så det ville jeg ha gjort da, i f.eks. python og algoritmer. Videoer, animasjoner, illustrasjoner osv.

I: Hva med å ta kontakt med enkeltstudenter du ser gjør det dårlig? Er det noe du også tror du ville ha gjort?

P11: Hvis jeg ser at en høy risk for å ikke bestå, vi snakker om stryk, så ville jeg nok ha tatt direkte kontakt med studenten. Men jeg ville nok ikke gjort det gjennom et slikt system. Da tenker jeg på videregående da. Tror ikke jeg hadde fått svar fra eleven hvis jeg hadde tatt kontakt gjennom systemet. Men jeg hadde brukt informasjonen fra dashboardet og agert på det utenfor systemet.

I: Skjønner. Er det noen funksjonalitet du savner?

P11: Hmm, nei, jeg tror ikke det.

I: Under brukertesten så nevnte du at det ikke var noen tilbakeknapp og noen småting.

P11: Det er jo litt teit av meg å si det da, for jeg vet jo at hvis jeg trykker på homeknappen så kommer jeg tilbake. Men jeg kunne gjerne tenkt meg å trykke videre på statistikken og fått opp flere av de dårligste spørsmålene. Jeg kan jo få tak i det fra question editoren, men det hadde vært greit å kunne ta tak i det direkte fra dashboardet.

I: Da har jeg bare ett spørsmål igjen. Ville du ha brukt SmartU applikasjonen hvis du hadde hatt muligheten, og hvor ofte ville du ha brukt den?

P11: Ja, jeg ville ha brukt den, både til prøver og løpende undervisning og vurdering. Hvor ofte jeg ville ha brukt den er avhengig av hvor mange timer jeg har i uka av et fag. Hvis jeg hadde hatt 3 timer naturfag i uka fordelt over 2 økter ville jeg nok brukt det 2 ganger i uka. Liker jeg å tro da.

Interview participant 12

I: Okay, so first of all, how did you feel about using the SmartU dashboard?

P12: I thought it looked pretty good. I felt confident that without any training I could just go in and click around and find out. If I was actually gonna use it for a class I would want more time, but I don't feel like I would need like a user manual or any hard instructions at all, so it seems pretty accessible.

I: Could you mention one or more functionalities or features that you felt were extra useful, and perhaps why?

P12: Knowing exactly which students exactly are struggling. So by having in the lists that Maren or whatever her name is had the little exclamation point on her so I knew right away. And then being able to see like what topics were the hardest for her and how she scored on those, and how she was doing relative to the rest of the class in the progress graph over time and over attempts. I think those things were super super relevant, so that I could have a good idea of where she's at scholastically before I reach out to her and see what ways I can support her if I need to do that.

I: Were there any features that you felt were confusing or difficult to understand?

P12: No. I wasn't a hundred percent sure exactly when I opened up and I clicked on a student and it opened up and showed me all of their attempts, and it was like nine out of nineteen, seven out of blah blah blah. I think I – actually there was one thing I didn't know how to do, I can't remember what it was but I had to click around to figure out how to do, and it had to do with wanting to reach out to a student or...

I: Yes it looked like you had to click around when you wanted to find their contact information.

P12: Yeah yeah yeah, and actually I don't even ever did, I think I just said I would contact her somehow. I figured out through the system, but I wanted to – yeah but that's what it was, I clicked on her and I wanted to bring up a student profile, but it brought up, where I clicked it brought up what all her tasks were over time. So I would like an easy or maybe more obvious way to click on a student and to be able to message them directly from where I am looking at the information. I don't want to have to go into a different spot because I want to be able to write that information while I'm looking at – like write her the email while I'm actually looking at her charts and what not compared to everyone.

I: As it was, it was only the email address, so the thought was "okay I will just use this email address", so I think my directions may have been a bit misleading there.

P12: Oh, like copy and paste that into – okay! No worries.

I: So out of the visualizations, which of the graphs did you prefer the most?

P12: Probably the progression over time and over attempts. Being able to see specific students as like a line chart thing. Being able to see a student relative to the class or the class average.

I: So on the other hand, were there any that you didn't like?

P12: No.

I: You did mention about contacting students, those types of features. Are there any other features that you would wish to see in the system?

P12: No, no I want to be able to contact them through while looking at their data, and not have to open up a separate thing.

I: Do you think you could utilize the information that was presented on the dashboard for something useful?

P12: Yeah, for like orchestrating future learning tasks based on all of the statistics about the questions. And also knowing what the distribution of the topics, the themes and the difficulty level across those were. By looking at those things you can see this is what I've laid out for the kids, and this is how they're doing and what questions are the hardest and potentially why. So by combining those and also knowing what the class average is looking like, I think that I could use that lots to kind of gear what the next content should be or if I should change the content that I have provided.

I: You have mentioned that you wanted to reach out to Maren (worst scoring on last attempt) in the system so I think we can skip that question. Do you think you could use the information to change the curriculum, maybe in the semester you are in or for next semester?

P12: Yeah for sure. I mean that kind of goes back to what I was just saying about looking at how like – how I laid out all the questions and the difficulty levels of the different themes and seeing how the students are doing and okay that worked, that didn't work, or this is not working right now and I need to change it right now. So I think it could be used for both present and future.

I: How did you feel about the dashboards design?

P12: I thought it was really good, I said that to you earlier.

I: Did you feel it had too many elements or too few?

P12: No, no there was definitely a lot going on but I think it was laid out in a way that I could manage looking at all of it at once, and I don't have to look at all of it at once. But some of those things you do want to look at, at the same time. So it's better to kind of have it all there, and just look at what you want instead of having to click back and forth between different tabs to see different view in my opinion.

I: How did you feel about the notification, the red circles?

P12: I think it's fine. They were a little small, but that might just be because I have old person eyes. But I thought they were fine, and the red one on the purple showed up a lot so yeah it

was fine. It told me like “there’s a problem here”. The only thing that I think may need to be revisited is the fact that zero logins was green, which I think zero logins for a system where the students that if they’re supposed to be using the system is bad, and green does not mean bad, green means good so I think that one needs to be flipped.

I: For this next question I think we can open up the view here. **opens up the filters on the mastery grid**. For the list here of students we have them sorted by their last quiz score, and an option is also to see the average score. How do you feel about score best represents a student, what would you like to see in a list like this?

P12: Oh sorry the way you have it is an average score?

I: Right now it shows the score from the last attempt, so that’s why [worst student] is at the top.

P12: So this is for her last – okay.

I: So if you use the filters to change from last attempt to average, you get different names.

P12: I missed that. I missed that.

I: So the question is basically how you feel about –

P12: Well now that I know that it’s here... ok so what’s the question?

I: Which kind of score do you feel is most representative, what would you like to see?

P12: I think having both is important because, ok so this is the last attempt so she sucked the hardest so why is that, but if I look at the average these are the kids that are actually really not doing the best. And so I missed that completely, but these kids need to be probably thought about as well.

I: But if you saw that they had improved and their last attempts are very good, but their first attempts are terrible, then maybe the average is not representative but for some it might be.

P12: Yeah and so I think it’s important to have both. You should be able to see how they did in the last because maybe it was just a one off and they did super well or they did super crappy. So I think the average is important because that’s like the trend for that student, but it’s important to know how they did on the last for sure.

I: One final question, how often do you think you would use an application such as SmartU?

P12: If I was a teacher? As long as my content was online, then I don’t see why I wouldn’t use it. I would use it. All the time. Depends on what my curriculum is, but I would use it as often as it was applicable to the content that I was serving to my students. I think it’s awesome.

Interview participant 13

I: Hva synes du om å bruke SmartU dashboardet?

P13: Jeg synes det virker bra. Fra det lille jeg så synes jeg det var greit.

I: Kan du nevne en eller flere funksjonaliteter som du syntes var ekstra nyttig og hvorfor?

P13: Ja, det er nyttig å se hvor mange som har klart oppgavene. Både for å se kvaliteten på oppgavene og for å se hvor godt undervisningen fungerer.

I: Var noen av funksjonalitetene forvirrende eller vanskelige å forstå?

P13: Ikke som jeg kommer på. Det kan hende det var noen funksjonalitet jeg ikke fant, men av det jeg så var ingenting forvirrende.

I: Hvilken visualisering i dashboardet likte du best, og hvorfor?

P13: Ja, det var veldig greit når man kommer inn til å begynne med å se de røde studentene på første side. For de grønne studentene er ikke så viktige siden de gjør det bra. Det er viktigere å se de studentene som ikke gjør det så bra.

I: Hvilken visualisering i dashboardet likte du minst, og hvorfor?

P13: Hmm... Nei, jeg kommer ikke på noe.

I: Var det noe funksjonalitet du skulle ønske var til stede på dashboardet?

P13: Jeg vet ikke om... Går det for eksempel an at man kan ha et prosjekt hvor elevene har en mulighet for å reservere seg mot at læreren kan se hva de har fått til?

I: Pr. i dag er det ikke det.

P13: Det kunne jeg tenkt meg å se. Hvis læreren oppdaget at det gjør studentene stresset og om de føler seg overvåket, så hadde det vært greit å ha en funksjonalitet for å kunne slå det av.

I: Tror du at du kunne ha brukt informasjonen som ble presentert på dashboardet til noe nyttig?

P13: Det tror jeg absolutt. Jeg kunne ha laget nye oppgaver, evaluert forelesning og undervisning.

I: Hvordan tror du at dashboardet kunne hjulpet deg med å nå ut til studenter som står i fare for å stryke?

P13: Ja, gitt at de bruker systemet, så tror jeg det.

I: Hvordan tror du at dashboardet kunne hjulpet deg med å gjøre passende endringer til læreplanen?

P13: Ja, det tror jeg, så lenge jeg legger arbeid i det og lager gode oppgaver så tror jeg det.

I: Hva synes du om dashboardets design?

P13: Jeg synes det virket veldig bra, pent og lett å bruke.

I: Ikke for overveldende?

P13: Nei, jeg synes det var akkurat passe.

I: Du har jo ikke fått så mye tid med systemet siden brukertesten er litt kort, så du fikk ikke sett så mye på de her røde notifikasjonene. Men jeg fikk med meg at du så den røde trekanten på studenten i mastery grid.

P13: Ja.

I: Hvis du hadde en lærer som bruker systemet her, tror du at en slik highlighting eller notifikasjon hadde hjulpet deg med å finne hvor du bør rette blikket ditt?

P13: Ja, men jeg vet ikke helt hvordan den er regnet ut.

I: Det går på gjennomsnitt.

P13: Ok, så gjennomsnitt. Det er jo nesten bedre å bare se på det røde. Hun som er under den øverste sliter jo kanskje mer som har 3 røde. Jeg synes kanskje fargekoden er det beste her, så den trekanten er kanskje ikke så nødvendig.

I: På høyresiden av dashboardet ser du en statistikk over noen av de spørsmålene som har flest feil, der det kommer en rød notifikasjon hvis feilraten er over en viss grense. Tror du det er greit med en sånn notifikasjon der?

P13: Ja, men det kunne ha vært sånn at hvis man klikker inn på det, så kunne det kommet opp hva som er nummer to på lista og sånn. Og den med longest time to answer er litt dum dersom en student starter et spørsmål og går fra PC-en.

I: Alle besvarelser over 10 minutter har vi filtrert ut, så de er ikke med i denne beregningen.

P13: Aah, det var lurt. Men hvis man har et vanskelig spørsmål i f.eks. fysikk så kan det faktisk ta 2 timer, eller iallfall en svartid på langt over 10 minutter.

I: Skjønner, da kunne det kanskje vært greit om læreren kan bestemme hva grensa er der.

P13: Ja, kunne ha fungert.

I: Jeg spør litt mer om de notifikasjonene jeg. Synes du det er forstyrrende at de er der, eller er det greit?

P13: Ja, jeg synes det er bra. Det er viktig at det ikke er for mye, men om det er 3 slike røde prikker så går det bra.

I: Når du da trykker på det spørsmålet med en notifikasjon så vil prikken forsvinne.

P13: Aah, så hvis jeg går inn og trykker cancel så forsvinner den? Skjønner. Enn hvis jeg trykker på Maren, forsvinner trekanten også?

I: Nei, den blir faktisk stående fram til hun gjør det bedre.

P13: Ja, men det er jo bra! Det kan jo hende jeg må ta kontakt med henne.

I: Så har jeg et spørsmål om måten score og studentenes risiko for å stryke er sortert på. Sånn som Maren som er på toppen her. Akkurat nå er det sortert på det siste forsøket. Så de med de dårligste siste forsøkene vil havne øverst. Vi har også lagt inn en filteringsmulighet for å filtrere på gjennomsnittlig score. Men vi er veldig usikre på hva som vil være mest nyttig for en lærer.

P13: Hvordan fungerer quizene. Velger de antall spørsmål, eller får de et bestemt antall spørsmål?

I: Antallet spørsmål og hvilke spørsmål er tilfeldig og vanskelighetsgraden tilpasses ut fra hvordan de har gjort det tidligere.

P13: Hvis det f.eks. bare er 10 spørsmål er det et veldig lite tallgrunnlag å bruke for å vurdere. Så siste forsøk høres ikke ut som en pålitelig måte å vurdere studentens nivå.

I: Godt poeng, kanskje om man tok en gjennomsnittlig score fra de siste 3 forsøkene eller noe kunne vært en løsning?

P13: Ja, jeg tenker at de tre siste høres ut som en god ide. Ellers så kunne man også 3 sorteringsmuligheter, for alle er ganske viktig. Så kan man jo velge selv. For både gjennomsnittlig og last attempt er greit å ha tilgjengelig.

I: Hvor ofte tror du at du ville brukt en applikasjon som SmartU?

P13: Da ville jeg sett hvordan det ser ut for studenten også. Jeg er i utgangspunktet litt skeptisk til å bruke IT i undervisning. Jeg synes ikke man skal bruke IT når det finnes gode alternative. Men det her er jo sånn som det ikke finnes noen gode alternativ til, som jeg har troa på. Så hvis jeg var en lærer så tror jeg at jeg ville ha gjerne tatt i bruk det her da. Det er også greit at jeg kunne ha laget spørsmål på forhånd, og brukt mer tid på retting av oppgaver. Det her retter jo automatisk, så det er jo veldig effektivt. Kunne jo ha laget det ferdig i sommerferien eller noe sånt. Så jeg tror jeg ville ha brukt det hver uke eller noe sånt.

C Design and implementation

C.1 User tests

Iteration 1 evaluation

1. Which course would you have selected, and why?
2. Explain what you see in the top three gauges of the dashboard.
3. For the bottom left statistics, please explain what you see.
4. What does the mastery grid tell you?
5. Where would you click to open up a student's details?
6. What does the "previous quizzes" tell you?
7. What does the line chart in the middle tell you?
8. Please explain the visualization to the right.
9. Was there anything in particular you liked/disliked?
10. Was there something you felt that was missing?

Q	User 1	User 2	User 3
1	Would select activity TDT31 because it has the red notification	Would select TDT31 because it has the red notification	Would select activity TDT31 because it has a red notification
2	Some measurement on web development topics, not sure what. Green is good, yellow medium and red is bad	Maybe how well students are doing in these topics, but not sure how it is calculated	Looks like some measurements on different topics. JavaScript seems to be the easiest topic
3	Deliveries of exercises in the course, but not sure	Number of people who did the quiz	Looks like 6 students have not tried the quiz yet
4	How each student is doing in each topic, same color coding as the gauges	Student list with their scores or performance in different parts of the course. The student at the top is highlighted probably because they are not doing very well	One of the students has two red topics, which can indicate that they are struggling. Can also see that the topic in the middle has more red dots compared to the other topics.
5	I can only see a student list in that mastery grid so I	Maybe there is a student list somewhere? Or is that	I would click on the student name in the

	guess I would click on a student there	the mastery grid?	mastery grid
6	Looks like scores on quizzes the student has done in the past in percentages	Previous scores on quizzes, maybe it is sorted from lowest to highest.	Looks like scores on previous quizzes.
7	A line chart showing score over time. I'm not sure what the difference is between the two lines, but one looks like a "ghost" line so maybe it is the best performing in the class for comparison?	A visualization of performance in the quizzes. The last three attempts match with the "previous quizzes" it looks like so the black line with colored dots is probably Ola Nordmann. Then maybe the blue is the average of the class.	Looks like a graph showing the score on the different attempts. Red is low score, yellow medium and green is good. Looks like the students has completed 4 attempts, but only 3 is showing in the previous quizzes. Not sure what the blue line means.
8	I don't understand why time is on the y-axis. Initially I thought it was the number of correct or wrong answers, but when I noticed time on the y-axis I suddenly have no idea what I'm looking at.	This one is difficult to understand, but maybe it is time spent on right and wrong answers. It is very unclear what this is actually showing.	The y-axis says time, and x-axis correct and wrong, so I assume it means how many correct and wrong answers over time?
9	I liked the mastery grid the most. The three gauges at the top and some of the student graphs were a bit difficult to interpret.	I liked that the colors were consistent, red was always negative and green was always positive and it correlated with the score or performance. The mastery grid was also okay. Some graphs were difficult to understand though.	I really liked the graph showing the attempts, and also the mastery grid. I don't think the "24/30 delivered" is that interesting.
10	I don't know specifically, but I think some more graphs could be helpful.	Some explanations on the graphs, maybe better labels on the axes to make them more understandable.	I would like to see some statistics on quiz questions. I also miss seeing an overview of the different questions.

Iteration 2 evaluation

1. *You are using the SmartU dashboard for teachers and have entered an activity. You are now in the dashboard view, which visualizes data from adaptive quizzes taken by students.*
Please explain what you are seeing in the graphs in the “Overview” card.
2. Explain what the visualization titled “Learning development” is showing.
3. Next you are viewing the “Mastery grid”. Please explain what you see.
4. Finally direct your attention to the “Activity insights”. What do you see here?
5. Where would you click to open up a student’s details?
6. What does the “previous quizzes” tell you?
7. Was there anything in particular you liked/disliked?

Q	User 1	User 2
1	Low, medium and high proficiency in percentage for students. I guess proficiency is the same as score so this is probably showing how many students do well, average and poorly.	I see an overview of students’ proficiency, divided into low, medium and high proficiency. Since the system uses adaptive quizzes, I guess this is about the quiz score.
2	There is the list and the graph. I see that the graph says “Showing for student 1” and it looks like the student is selected in the list. It looks like there are two lines, and since the gray-blue one matches the color that is around the student name I guess that is the students development over time, while the multi-color line is perhaps the average or median for the rest of the class. Looks like Student 1 is close to that line so that is probably a good thing.	I can see a student list and a graph. The top student is highlighted, so I assume the student is selected and the graph is showing the student’s score from week to week. The colors are confusing, as there are two lines, one with some red, yellow and green, while the other one is blue. I don’t know which line is showing what.
3	I see that this has the same students as in the previous graph. But I notice student 1 is not doing so well here if I am understanding this correctly. Horizontally it says concepts, I guess those are learning goals or something similar. So student 1 is doing poorly in all three concepts. Then there is a	I can see a student list and a list of three concepts. I assume red is bad and green is good. The student on the top is red in all concepts and is also highlighted with a yellow triangle.

	small warning sign there which probably means I should take a look at that.	
4	I see two questions. One is most commonly failed, so that's probably a difficult question. The other is the longest time to answer, so maybe that one is also difficult.	I can see the most commonly failed question and the question with the longest time to answer.
5	There are two lists of names but it looks like the upper right one is used to change the graph so probably the one in the mastery grid list.	I would click on one of the two student lists. Maybe the mastery grid since it has the triangle with the exclamation mark.
6	Percentage of correct or score and which date it was done. So it looks like the latest date is at the top.	I guess it tells me the percentage of correct answers on previous attempts.
7	The activity insights with the most commonly failed and longest time to answer looked a bit messy with all the text all over the place, but I think they could be really helpful to find problems in the quiz. I also think that it could be hard to find struggling students in the learning development graph in an activity with a large user base.	I really like the learning development graph, but I would like an explanation of the two different lines. I mean the difference between the blue and multi-colored lines.

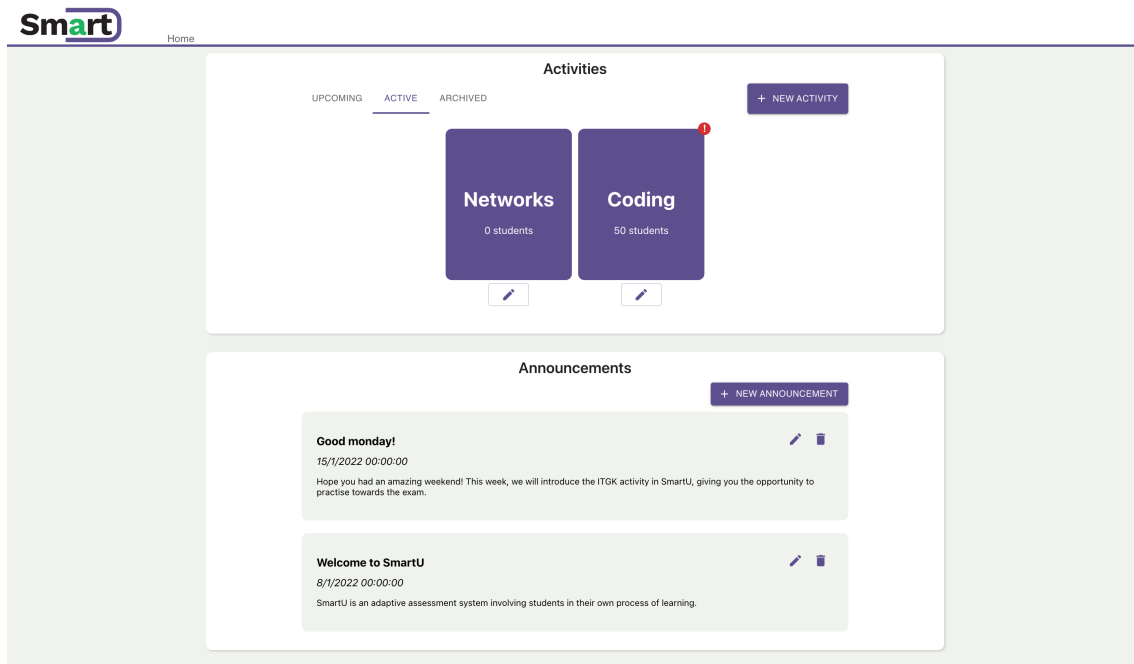
Iteration 3 evaluation

1. *You are using the SmartU dashboard for teachers and have entered an activity. You are now in the dashboard view, which visualizes data from adaptive quizzes taken by students. You are already logged in, and see the landing page.*
Which activity would you have selected, and why?
2. What do you think about the announcements?
3. Please explain what you are seeing in the graphs in the “Overview” card.
4. Explain what the visualization titled “Learning development” is showing.
5. Next you are viewing the “Mastery grid”. Please explain what you see.
6. Finally direct your attention to the “Activity insights”. What do you see here?
7. Was there anything in particular you liked/disliked?

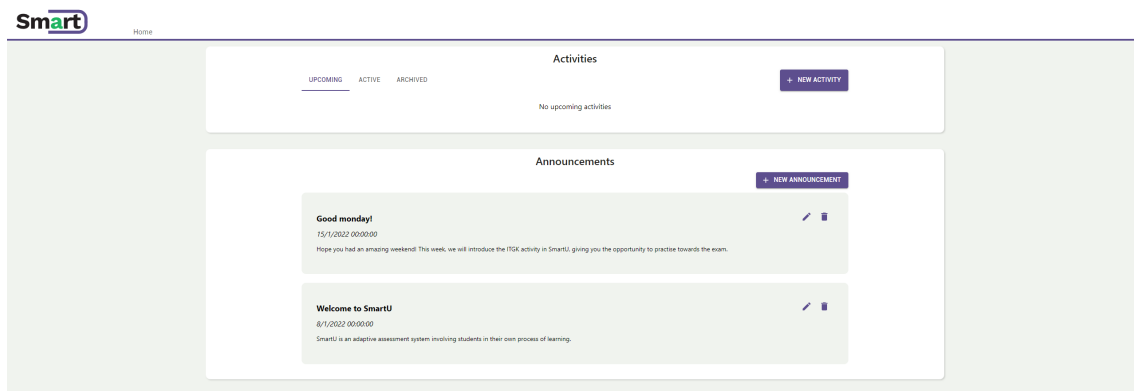
Q	User 1	User 2
1	I see three activities here, but the one on the right has some kind of notification on it so I would click on that one.	I see that activity TDT4080 has a red notification, so I think I would click on that one.
2	Looks alright, pretty straightforward that it shows announcements. But if I am a teacher where do I press to create one, if it is intended that I can do that?	The announcements looks ok. Is it possible to post announcements for only one activity?
3	This looks like a distribution of student proficiency represented by a bar chart, although I think a pie-chart would be a better chart for this purpose.	I can see the percentage of students with low, medium or high proficiency.
4	How students are developing over time. On the x-axis we have weeks, and the y-axis shows score. On the explanation on the right I see that the purple line is the class average and that the yellow is the student selected from the list on the left. The yellow line is much lower than the class average so this person is not doing very well. The colors on the names in the list on the left looks to be	I can see a list of students. The two students on the top are red, the next two students are yellow and the rest are green. I assume this means the students with low proficiency is showing at the top. This is nice. I would like to have more sorting options. I can also see the class average score compared to individual students. The coloring on the graph is a bit confusing,

	color coded so it makes sense since the selected student is red in the list.	as it seems like a red student is selected, but the line is yellow.
5	Given the title I suppose this is showing how each person listed has mastered the different topics. This is also using the same color scheme as before so it is easy to see that red is negative.	Looks like this visualization is sorted the same way, with the lowest proficiency at the top. I can see which topics each student has mastered, which is nice.
6	Two types of statistics. The first one shows which question that has been answered incorrectly the most, with a large percentage of answers being wrong. The second shows which question with most time spent on. Not sure if that is the longest time from one answer or as an average.	I can see some statistics on two different types of questions: most commonly failed question and longest time spent on. I can see the usefulness of knowing which question has most wrong answers, but the longest time spent is a bit confusing. What if a question is long and should take some time to answer?
7	I liked most of it, it was easy to understand. My only problem was the placeholder names which made it a bit difficult to understand if the students in the learning development graph and the students in the mastery grid were the same people and if they were sorted the same way. I liked the statistics too, and I think there is possibly more cool statistics that could be shown that would be of interest. As long as you don't have to scroll to see more statistics that is.	As I mentioned before, I would like to have more sorting options, such as alphabetical sorting and ascending/descending order.

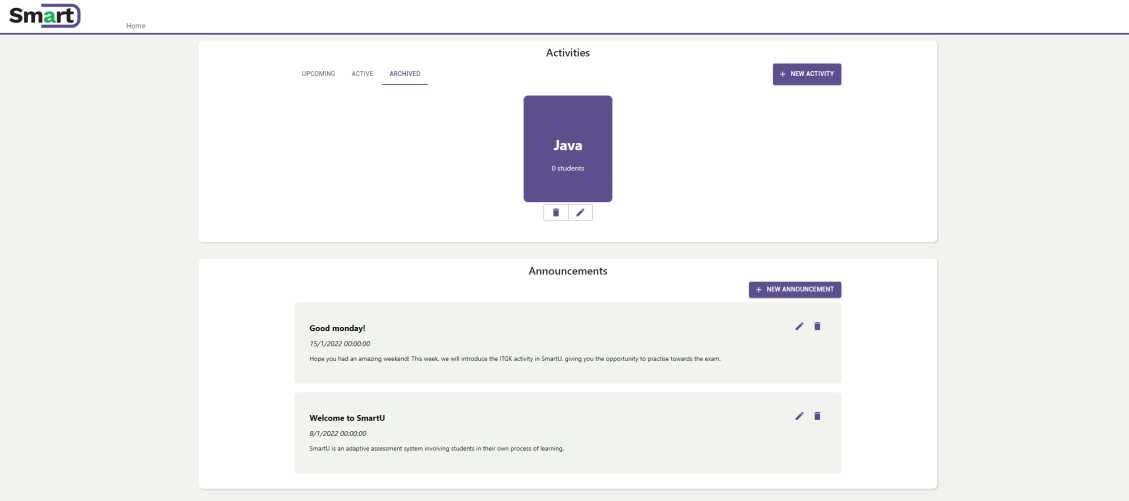
C.2 Final user interface



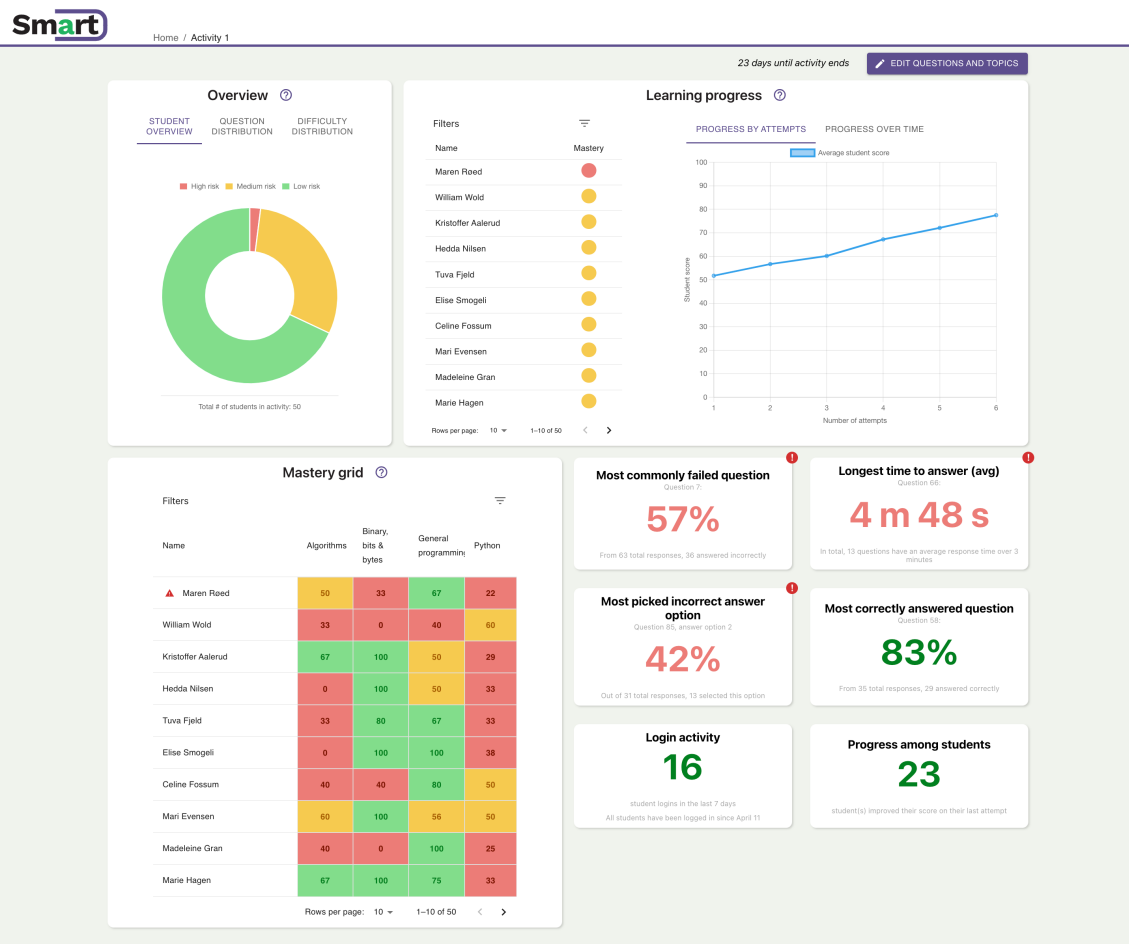
Application frontpage



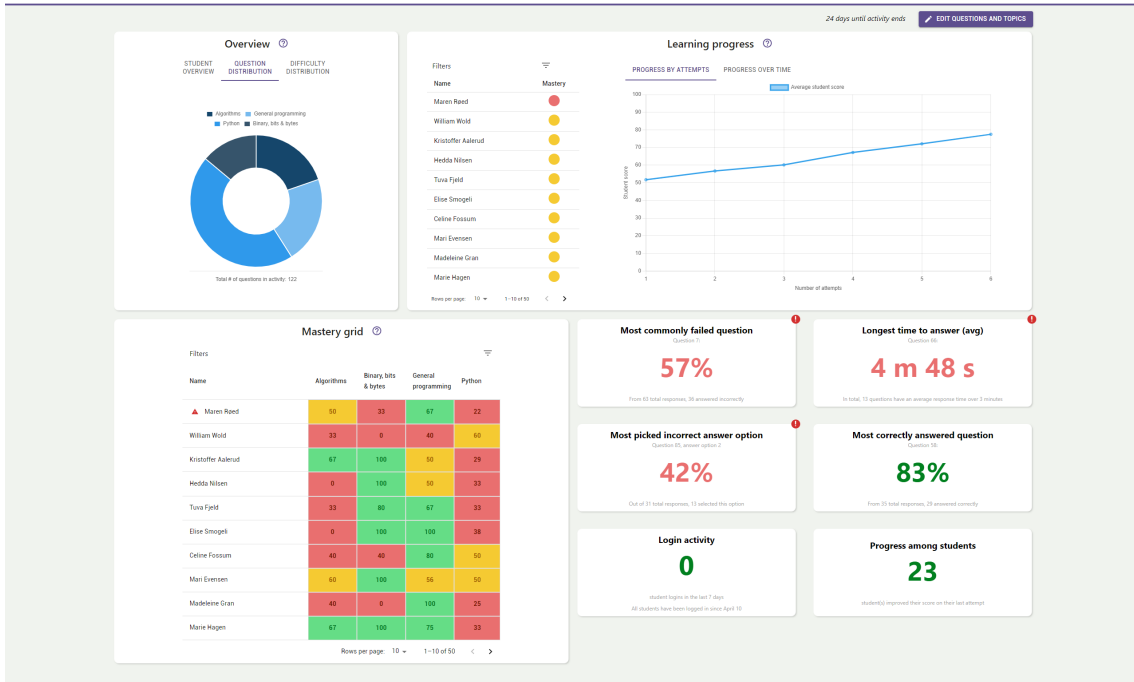
Application frontpage upcoming activities tab



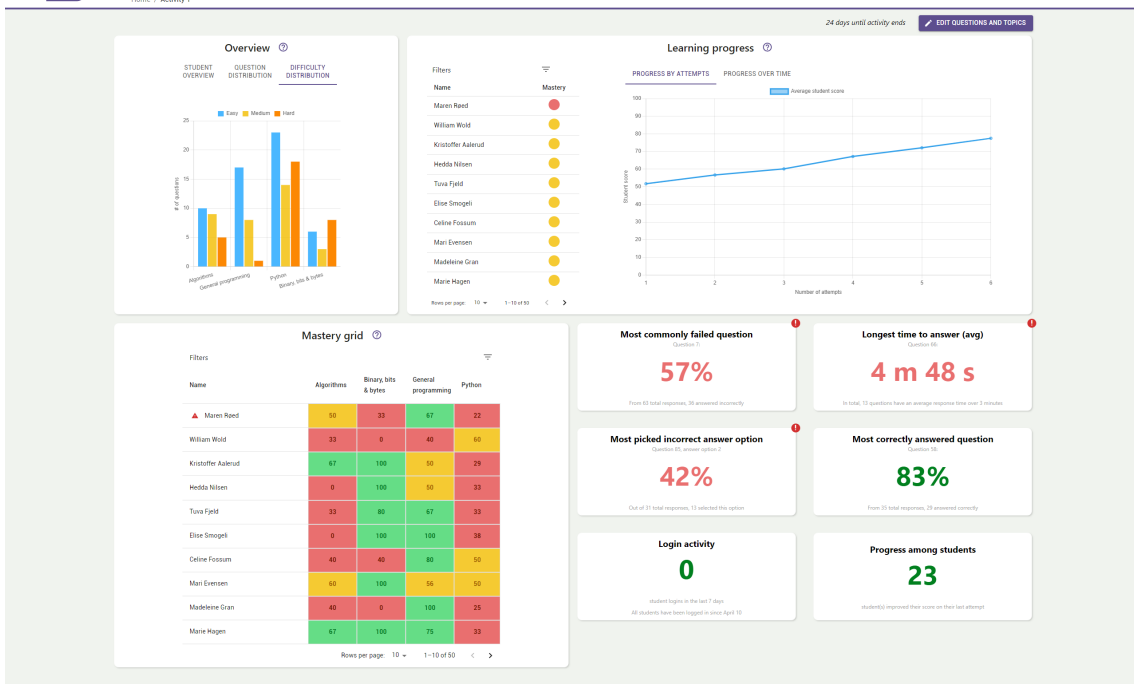
Application frontpage archived activities tab



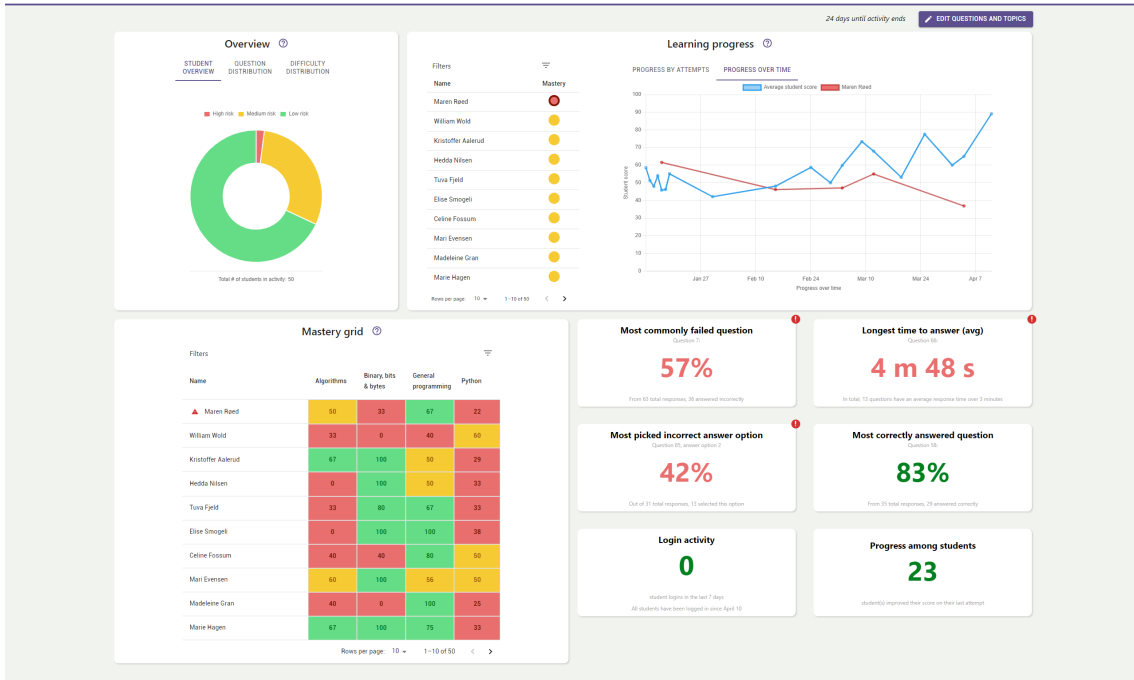
Dashboard main view



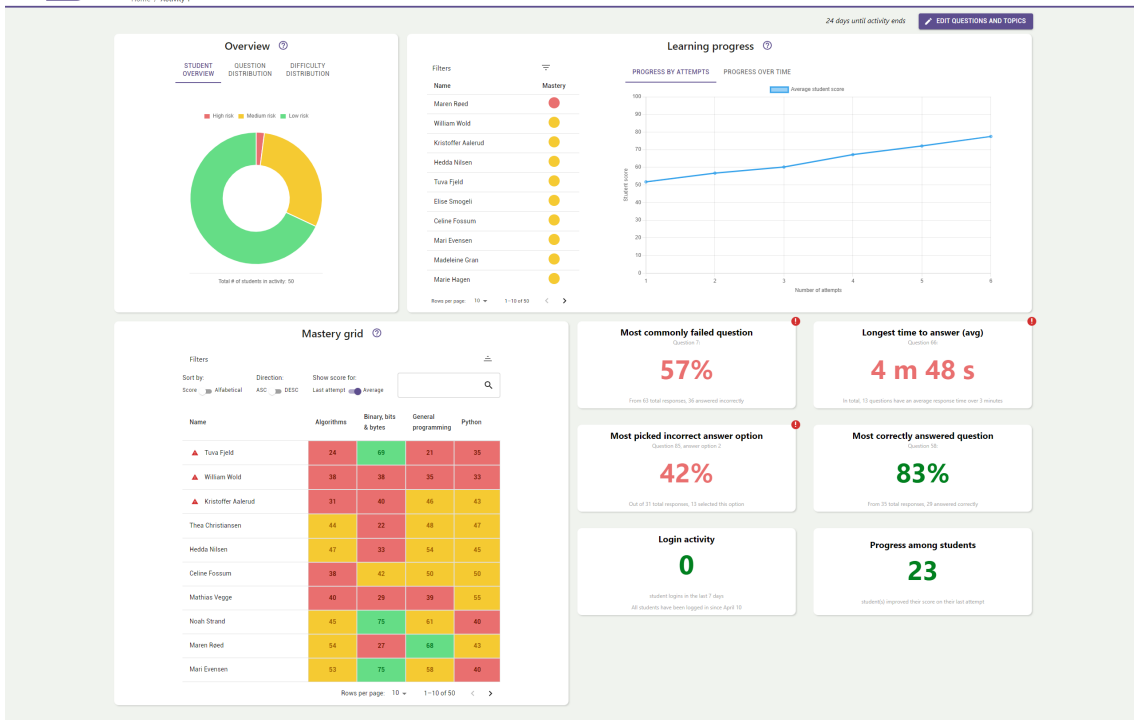
Dashboard overview question distribution



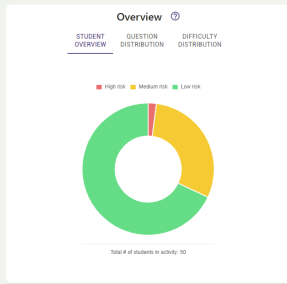
Dashboard overview difficulty distribution



Dashboard learning progress graph with student selected



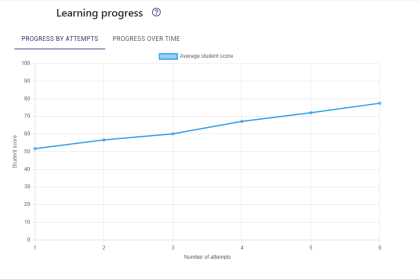
Dashboard mastery grid sort by score



Filters

Name	Mastery
Maren Raed	●
William Wold	●
Kristoffer Aabrud	●
Hedda Nilsen	●
Tova Fjell	●
Else Smogeli	●
Celine Fossum	●
Mar Eivensen	●
Madeleine Gran	●
Mare Hagen	●

Rows per page: 10 1-11 of 50



Mastery grid

Filters

Sort by: score Alphabetical ASC DESC Last attempt Average

Show score for:

Name	Algorithms	Binary bits & bytes	General programming	Python
Tova Fjell	33	80	67	33
Tova Finstad	100	100	100	73

Rows per page: 10 1-10 of 50

Most commonly failed question

Question 7

57%

From 63 total responses, 36 answered incorrectly

Longest time to answer (avg)

Question 66

4 m 48 s

In total, 10 questions have an average response time over 3 minutes

Most picked incorrect answer option

Question 66

42%

Out of 31 total responses, 13 selected this option

Most correctly answered question

Question 66

83%

From 35 total responses, 29 answered correctly

Login activity

0

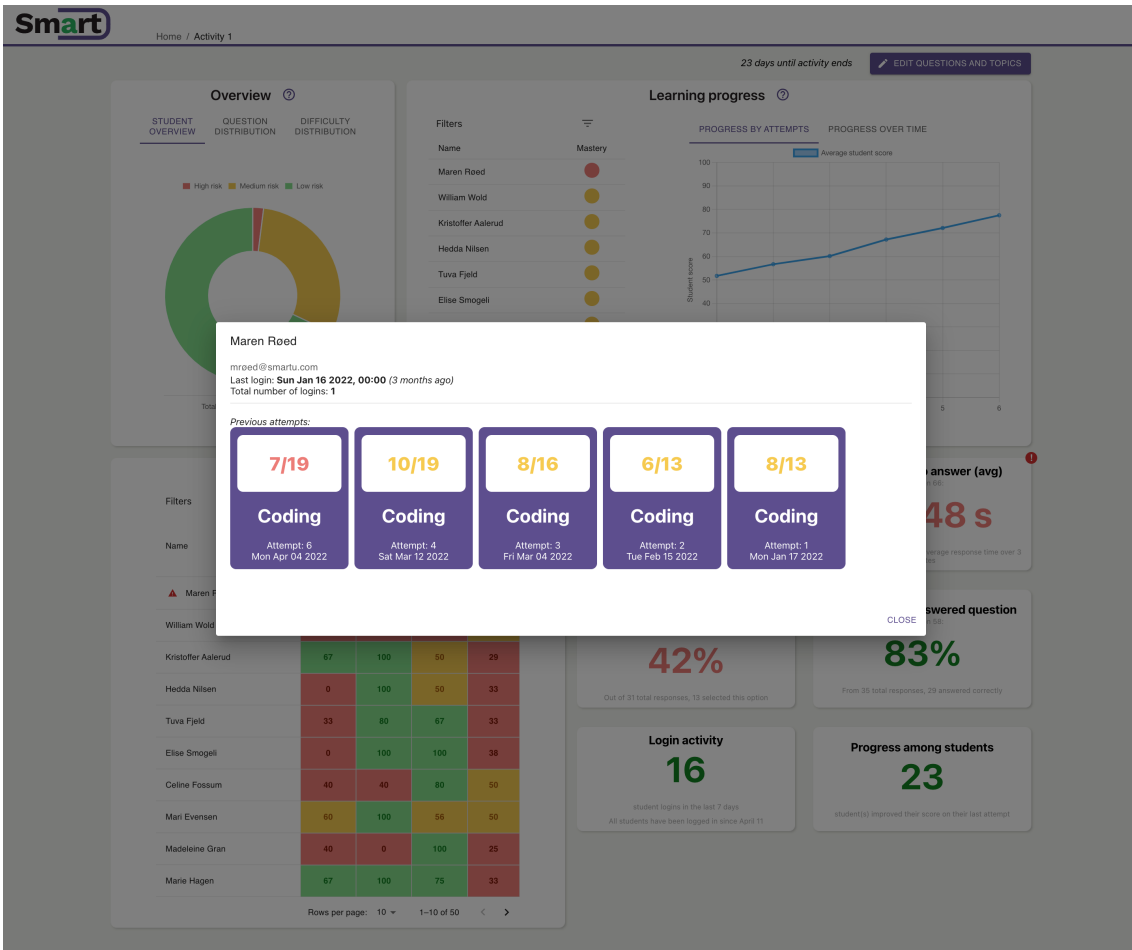
student logins in the last 7 days
46 students have been logged in since April 10

Progress among students

23

student(s) improved their score on this test attempt

Dashboard mastery grid search



Dashboard mastery grid student details

The screenshot displays a dashboard for a course activity. At the top, it says 'Smart' and 'Home / Activity 1'. A notification indicates '23 days until activity ends' and a button for 'EDIT QUESTIONS AND TOPICS'. The main area is divided into 'Overview' and 'Learning progress'. The 'Overview' section includes a donut chart for 'STUDENT OVERVIEW' (High, Medium, Low risk) and a 'Filters' table. The 'Learning progress' section features a line graph for 'PROGRESS BY ATTEMPTS' and 'PROGRESS OVER TIME'. A modal window for 'Maren Roed' is open, showing 'Attempt 6' with a breakdown of scores: Algorithms: 2/4 (yellow), Binary, bits & bytes: 1/3 (red), General programming: 2/3 (green), and Python: 2/9 (red). The background shows a mastery grid for various students and several summary cards: '42%' (13 answered), '83%' (29 answered), 'Login activity 16' (16 logged in), and 'Progress among students 23' (23 improved).

Dashboard mastery grid student details after clicking on an attempt

Smart Home / Activity 1 23 days until activity ends EDIT QUESTIONS AND TOPICS

Overview

STUDENT OVERVIEW QUESTION DISTRIBUTION DIFFICULTY DISTRIBUTION

High risk Medium risk Low risk

Total # of students in activity: 50

Learning progress

Filters Name Mastery

PROGRESS BY ATTEMPTS PROGRESS OVER TIME

Average student score

Number of attempts

Edit Question

Edit questions, change answer options, change correct answer etc.

Question*

Definisjon av algoritme er i følge teoriboka et ordnet sett av ...

Velg fi Ingen fil valgt

(Optional) Upload an image

Option 1* Set correct

enlydige, utforbare steg som def en terminerende prosess

Option 2*

tvetydige steg som løser et gitt problem effektivt

Option 3*

alger som spiller slagverk eller andre rytmeinstrumenter

Option 4*

sekvenser som utføres enten i parallell eller sekvensielt

+ ADD OPTION

Topic

Algorithms

Select a topic for the question

Difficulty

Hard

Select a difficulty rating for the question

CANCEL SAVE

Mastery grid

Filters

Name	Algorithms	Binary bits & bytes		
Maren Raed	50	23		
William Wold	33	0		
Kristoffer Aalerud	67	100		
Hedda Nilson	0	100		
Tuva Fjeld	33	60		
Elise Smogeli	0	100		
Celine Fossum	40	40		
Mari Evensen	60	100	56	50
Madeleine Gran	40	0	100	25
Marie Hagen	67	100	75	33

Rows per page: 10 1-10 of 50

Longest time to answer (avg)

4 m 48 s

In total, 13 questions have an average response time over 3 minutes

Most correctly answered question

83%

From 35 total responses, 29 answered correctly

Progress among students

23

student(s) improved their score on their last attempt

Student logins in the last 7 days: All students have been logged in since April 11

Dashboard highlighted statistics opened question 1

Overview

STUDENT OVERVIEW QUESTION DISTRIBUTION DIFFICULTY DISTRIBUTION

High 100% Medium 100% Low 100%

Total # of students in activity: 10

Mastery grid

Name	Algorithms	Binary bits & bytes	General programming
Maren Raed	0	33	67
William Word	33	0	67
Kristoffer Aslerud	67	100	0
Hedda Nilsen	0	100	0
Tora Fjeld	33	67	0
Elise Smøgtli	0	100	100
Celine Fossum	67	67	0
Mar Evensen	67	100	0
Maddeline Gran	67	0	100
Marie Hagen	67	100	0

Learning progress

PROGRESS BY ATTEMPTS PROGRESS OVER TIME

24 days until activity ends EDIT QUESTIONS AND TOPICS

Edit Question

Edit questions, change answer options, change correct answer etc.

presidents = ['Barack', 'George', 'Bill', 'Franklin', 'Jimmy', 'Theodore', 'Richard']

Question: Hva må skrives i Python for å hente ut 'George', 'Ronald' og 'Gerald' fra lista presidents?

Vilg til Ingen til valgt

Optional: upload an image

Set correct

Option 1: presidents[1 : 2] eller presidents[1 : 6 : 2]

Option 2: presidents[2 : 4 : 4]

Option 3: presidents[1 : 3 : 5]

+ ADD OPTION

Topic: Python

Selected a topic for the question

Difficulty: Medium

Select a difficulty rating for the question

CANCEL SAVE

Longest time to answer (avg)
4 m 48 s

Most correctly answered question
83%

Progress among students
23

Dashboard highlighted statistics opened question 2

Home / Activity 2

236 days until activity ends EDIT QUESTIONS AND TOPICS

There is no data to display yet for this activity

Data will be displayed here once it exists

Empty dashboard with no data

Double click a question to edit

[TOPICS](#) [UPLOAD](#) [+ ADD](#) [DELETE](#)

ID	Question	Topic	Difficulty	Correct answers	Wrong answers	Effort	Performance	Average response time
1	Algoritmer er alltid organisert slik at man først	Algorithms	3	48	26	0.946	0.649	2 m 22 s
7	Definisjon av algoritme er i følge teoriboka et i	Algorithms	3	27	36	0.921	0.429	2 m 25 s
8	En algoritme kan beskrives som	Algorithms	1	37	19	0.946	0.661	1 m 56 s
21	Hva beskriver store-theta (Θ) notasjon i forbin	Algorithms	2	30	17	0.894	0.638	2 m 15 s
23	Hva betyr det at en algoritme skal ha en termi	Algorithms	1	27	20	0.894	0.574	2 m 21 s
53	Hva er tidskompleksiteten til denne pseudoko	Algorithms	2	25	19	0.932	0.568	2 m 31 s
54	Hva er tidskompleksiteten til denne pseudoko	Algorithms	2	35	14	0.959	0.714	1 m 52 s
55	Hva gjør denne algoritmen?	Algorithms	1	26	11	0.946	0.703	1 m 8 s
60	Hva kan du si om denne algoritmen?	Algorithms	1	27	6	1	0.818	2 m 25 s
70	Hva stemmer for den naive bondesjakkalgorit	Algorithms	1	22	11	0.97	0.667	1 m 41 s
71	Hva stemmer for den smarte bondesjakkalgor	Algorithms	2	19	15	0.941	0.559	2 m 27 s
72	Hva tidskompleksiteten til en algoritme der m	Algorithms	3	15	11	0.962	0.577	1 m 60 s
73	Hva tidskompleksiteten til program med enkel	Algorithms	2	18	18	0.889	0.5	2 m 3 s
74	Hva tidskompleksiteten til rent sekvensiell pro	Algorithms	3	19	14	0.939	0.576	1 m 46 s
77	Hvilke egenskaper har binærseek algoritmen?	Algorithms	1	28	14	0.952	0.667	2 m 10 s
78	Hvilke egenskaper har innstikksorteringsalgor	Algorithms	1	19	6	0.92	0.76	1 m 19 s
79	Hvilke egenskaper har sekvensiell søk?	Algorithms	2	18	13	0.903	0.581	2 m 21 s
82	Hvilken algoritme er dette?	Algorithms	1	20	17	0.892	0.541	1 m 20 s
83	Hvilken algoritme med følgende karakteristikk	Algorithms	3	17	12	0.931	0.586	1 m 30 s
84	Hvilken algoritme med følgende karakteristikk	Algorithms	2	21	9	0.9	0.7	1 m 55 s
85	Hvilken av disse algoritmene er mest effektiv i	Algorithms	2	18	13	0.935	0.581	1 m 59 s
107	Hvordan beskrives ofte algoritmer?	Algorithms	2	16	13	0.931	0.552	2 m 5 s
115	Må stegene i en algoritme være ordnet (følge	Algorithms	1	21	9	1	0.7	1 m 60 s
116	Med at en algoritme skal ha entydige steg me	Algorithms	1	24	12	0.889	0.667	2 m 16 s
2	Bestem True eller False: (A>5 or B==2)	General programming	1	44	19	0.984	0.698	1 m 46 s

Rows per page: 25 1-25 of 122 < >

[TO ACTIVITY DASHBOARD](#)

Edit question page

Smart Home / Activity 1 / Edit

Double click a question to edit

TOPICS UPLOAD + ADD DELETE

ID	Question	Topic	Difficulty	Correct answers	Wrong answers	Effort	Performance	Average response time
1	Algoritmer er altid organiseret slik at man først	Algorithms	3	48	26	0.946	0.649	2 m 22 s
7	Definisjon av algoritme er i følge teorboka et	Algorithms	3	27	36	0.921	0.429	2 m 25 s
8	En algoritme kan beskrives som	Algorithms	1	37	19	0.946	0.661	1 m 56 s
21	Hva beskriver store-theta (Θ) notasjonen i forbi	Algorithms	2	30	17	0.894	0.638	2 m 15 s
23	Hva betyr det at en algoritme skal ha en termi	Algorithms	1					2 m 21 s
53	Hva er tidkomplekseteten til denne pseudok	Algorithms	2					2 m 31 s
54	Hva er tidkomplekseteten til denne pseudok	Algorithms	2					1 m 52 s
55	Hva går denne algoritmen?	Algorithms	1					1 m 8 s
60	Hva kan du si om denne algoritmen?	Algorithms	1					2 m 25 s
70	Hva stemmer for den naive bondesjakkalgori	Algorithms	1					1 m 41 s
71	Hva stemmer for den smarte bondesjakkalgori	Algorithms	2					2 m 27 s
72	Hva tidkomplekseteten til en algoritme der m	Algorithms	3					1 m 60 s
73	Hva tidkomplekseteten til program med enke	Algorithms	2					2 m 3 s
74	Hva tidkomplekseteten til rent rekursivsett pr	Algorithms	3					1 m 46 s
77	Hvilke egenskaper har binærseek algoritmen?	Algorithms	1					2 m 10 s
78	Hvilke egenskaper har rekursivsettingsalgori	Algorithms	1					1 m 19 s
79	Hvilke egenskaper har selvsettsett sak?	Algorithms	2					2 m 21 s
82	Hvilken algoritme er dette?	Algorithms	1					1 m 20 s
83	Hvilken algoritme med følgende karakterist	Algorithms	3					1 m 30 s
84	Hvilken algoritme med følgende karakterist	Algorithms	2					1 m 55 s
85	Hvilken av disse algoritmene er mest effektiv	Algorithms	2					1 m 59 s
107	Hvordan beskrives ofte algoritmer?	Algorithms	2	16	13	0.931	0.552	2 m 5 s
115	Må stegene i en algoritme være ordnet i følge	Algorithms	1	21	9	1	0.7	1 m 60 s
116	Med et en algoritme skal ha enlydige steg me	Algorithms	1	24	12	0.889	0.667	2 m 16 s
2	Bestem True eller False: (A-5 or B+2)	General programming	1	44	19	0.984	0.698	1 m 46 s

Rows per page: 25 1-25 of 122

TO ACTIVITY DASHBOARD

Add question

Edit questions, change answer options, change correct answer etc.

Question +

Velg til Ingen fil valgt
(Optional) Upload an image

Option 1 +

Option 2 +

+ ADD OPTION

Topic: **No category**

Select a topic for the question

Difficulty: **Easy**

Select a difficulty rating for the question

CANCEL SAVE

Edit question page, add questions

Smart Home / Activity 1 / Edit

Double click a question to edit

TOPICS UPLOAD + ADD DELETE

ID	Question	Topic	Difficulty	Correct answers	Wrong answers	Effort	Performance	Average response time
1	Algoritmer er alltid organisert slik at man først	Algorithms	3	48	26	0.946	0.649	2 m 22 s
7	Definisjon av algoritme er i følge teorboka et	Algorithms	3	27	36	0.921	0.429	2 m 25 s
8	En algoritme kan beskrives som	Algorithms	1	37	19	0.946	0.661	1 m 56 s
21	Hva beskriver store-theta (Θ) notasjonen i forbi	Algorithms	2	30	17	0.894	0.638	2 m 15 s
23	Hva betyr det at en algoritme skal ha en termi	Algorithms	1	27	20	0.894	0.574	2 m 21 s
53	Hva er tidkomplekseteten til denne pseudok	Algorithms	2	26	18	0.922	0.568	2 m 31 s
54	Hva er tidkomplekseteten til denne pseudok	Algorithms	2					1 m 52 s
55	Hva går denne algoritmen?	Algorithms	1					1 m 8 s
60	Hva kan du si om denne algoritmen?	Algorithms	1					2 m 25 s
70	Hva stemmer for den naive bondesjakkalgori	Algorithms	1					1 m 41 s

Rows per page: 10 1-10 of 122

TO ACTIVITY DASHBOARD

Upload questions from file

You can upload files with .csv, .xlsx or .xlsx extension. You can also download a template with the desired format to get you started on making your own question file.

Select file format:

CSV

UPLOAD FROM FILE DOWNLOAD TEMPLATE

Edit question page, upload questions

Home / Activity 1 / Edit

Double click a question to edit

ID	Question	Topic	Difficulty	Correct answers	Wrong answers	Effort	Performance	Average response time
1	Algoritmer er alltid organisert slik at man først	Algorithms	3	48	26	0.946	0.649	2 m 22 s
7	Definisjon av algoritme er i følge teorboka et	Algorithms	3	27	36	0.921	0.429	2 m 29 s
8	En algoritme kan beskrives som	Algorithms	1	37	19	0.948	0.661	1 m 56 s
21	Hva beskriver store-Beta (B) notasjonen i forbi	Algorithms	2	30	17	0.894	0.638	2 m 15 s
23	Hva betyr det at en algoritme skal ha en termi	Algorithms	1	27	20	0.894	0.574	2 m 21 s
53	Hva er tidskomplekksiteten til denne pseudok	Algorithms	2	25			0.568	2 m 31 s
54	Hva er tidskomplekksiteten til denne pseudok	Algorithms	2	35			0.714	1 m 52 s
55	Hva gjør denne algoritmen?	Algorithms	1	26			0.793	1 m 8 s
60	Hva kan du si om denne algoritmen?	Algorithms	1	27			0.818	2 m 25 s
70	Hva stemmer for den naive bondetjakkalgorit	Algorithms	1	22			0.667	1 m 41 s

Rows per page: 10 | 1-10 of 122

TO ACTIVITY DASHBOARD

Edit question page, manage topics

Home / New Activity

Create a new activity

Title *

Minimum 2 characters, 18 characters left

Activity text *

Start date: 05/10/2022

Start time: 11:05

End date: 05/11/2022

End time: 11:05

CANCEL SAVE AND EXIT NEXT

Create activity

Edit activity

Title *

Coding

Minimum 2 characters, 4 characters left

Activity text *

This activity will let you test yourself on questions in Python, Algorithms, Binary and some general programming questions.

Archive activity

Start date
03/29/2022

Start time
11:02

End date
06/03/2022

End time
01:00

CANCEL SAVE AND EXIT NEXT

Edit activity

Create a new announcement

Title *

Announcement text *

CANCEL PUBLISH

Create announcement

Create a new announcement

Title *
Good monday!

Announcement text *
Hope you had an amazing weekend! This week, we will introduce the TGC activity in SmartU, giving you the opportunity to practise towards the exam.

CANCEL PUBLISH

Edit announcement

C.3 Code repository

The code for the project is available at: <https://github.com/fredrmon/smartu>

