# Continuing work on MarkEd: the tool for marking, moderation and feedback

*Leanora Andrusz*


4th Year Project Report
Computer Science
School of Informatics
University of Edinburgh

2024

# Abstract

This project centres on the development of MarkEd, a marking and feedback tool. Specifically, the project aimed to identify all remaining gaps following previous work on MarkEd and to implement the highest priority requirements in order to improve the tools's overall usability and impact. In order to achieve this goal, the work of prior students was closely analysed to reveal the requirements they had designed or partially implemented, together with user feedback. This formed the basis of an extensive list of features that still needed to be implemented. Using an adapted ICE prioritisation framework [17], every requirement in the list was scored and placed into one of three priority groups: "high", "medium" and "low". All requirements in the "high" priority group as well as some "medium" priority requirements were subsequently implemented using the Python web development framework Django [28], the CSS framework Bootstrap [13] and MySQL [69] database management system. Initially, the project was transferred from a Windows to a Mac machine to enable local development. After development finished, it was deployed on a Linux Debian server to enable user studies to be carried out remotely and asynchronously, with users testing the system on their own personal computers. All target user groups were involved in giving feedback on the usability and potential impact of MarkEd. This was achieved via a stakeholder walkthrough with academics and think aloud studies and interviews with markers, with all sessions being followed by a short questionnaire. A longer, more detailed questionnaire was sent out to students. The total 31 participants gave SUS [15] scores ranging from "good" to "excellent", with the system receiving scores around 4/5 as having potential for "ensuring fairness" in marking by teaching staff. The project collected a wide range of qualitative data, and overall the studies demonstrated that the system has great potential for improving marking processes at the School of Informatics.

# Research Ethics Approval

This project obtained approval from the Informatics Research Ethics committee.
Ethics application number: 681259
Date when approval was obtained: 2023-12-06
The relevant participants' information sheets and consent forms are included in the appendix.

# Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Leanora Andrusz*)

# Acknowledgements

I'd like to thank my supervisor Dr Alexandru, for the consistent help and guidance throughout this project. Of course I need to thank my parents at this stage as well, for their continued support and financing of coffees during this time. Shoutout to my little sister, who somehow always manages to put everything into the right perspective.

# Table of Contents

# Chapter 1

# Introduction

## 1.1 Motivation

Online coursework submission has become commonplace in Higher Education (HE) over the past 10 years, in contrast to the former practice of paper hand-ins [18]. This brings with it a variety of benefits such as allowing students to make a submission or read feedback remotely, providing a centralised place for markers to enter feedback and upload marks, and reducing paper consumption [86]. Additionally, online systems such as Blackboard Learn [56] often enable the integration of other tools such as plagiarism checkers.

The University of Edinburgh School of Informatics currently makes use of a number of online platforms to manage assessments:

- Blackboard Learn Ultra [56], the most widely used learning management system globally [86], for overall course materials, calendars, and marking tools.
- Turnitin [104], to check pdf submissions for plagiarism.
- Gradescope [40], which enables both code and pdf submissions, and contains the option of using an auto grader or AI-enhanced grading.
- CodeGrade [24], which was developed exclusively for marking coding projects and which allows for autograding and annotating code.

While these are some of the most popular platforms in the market at present [9], none of these systems provide the comprehensive set of tools that is needed for informatics courses [8]. Computer science courses have a high number of code projects that come in various formats, as well as essays and reports. This makes a run-of-the-mill system that only accept pdfs or zip files unsuitable for marking and at the same time providing feedback. To mark coding assignments, academics therefore often have to develop their own "auto graders" to perform automated checks on the code, using up valuable time not only to develop these tools but also then to export and import data back onto the main platform. There are also many other aspects of code assignments that require detailed inspection by the marker, including syntax, semantics, style (readability) and performance [6].

Additionally, the University of Edinburgh currently has some of the lowest scores in the National Student Survey in the UK [102], specifically regarding delivering marks on time, as well as delivering high-quality feedback [42]. A large reason for this can be attributed to marker numbers not keeping up with cohort numbers and therefore struggling to uphold deadlines [39].

MarkEd was developed for all the reasons above. It aims to provide the flexibility needed in a wide variety of assessment and marking formats, and at the same time incorporate strategies developed from existing research to ensure fair and efficient marking as well as high-quality feedback.

Examples of such strategies are shuffling submissions to avoid bias that can arise when submissions are always viewed in a certain order, or timed marking to help with a marker's personal time management and increase efficiency [23]. Indeed, research shows that communication between markers can help create higher quality feedback [111]. Many more strategies have been incorporated into the design, with preliminary user studies demonstrating potential for target users. Now in its fourth year of development, MarkEd can help to streamline the marking process and ultimately assist teaching staff at the School of Informatics in dealing with large volumes of assessments under tight deadlines, without the need to compromise on the quality or fairness of the marks.

## 1.2 Aims and research questions

The main aim of this project is to continue implementation and improve the overall usability of the system started by previous students. This will be achieved by identifying and implementing high-priority requirements and subsequently evaluating these changes with end-users. The desired outcome would be to observe an increase in perceived usability across user groups and simultaneously, demonstrating that the new features are a useful addition to the submission process for students and the marking process for teaching staff.

The research questions below will help guide the process and enable a deeper understanding of the current system and design, and how to best improve it over the project timeline.

1. What features are missing from the current system?
    (a) Based on feedback on the last students' work
    (b) Based on the original requirements and designs in the first three students' dissertations
2. Which features should be prioritised?
3. What additional steps need to be taken in order to:
    (a) Transfer the project onto a new local machine
    (b) Host the system online for user testing
4. How are the new changes perceived by the target user groups:

(a) students in terms of usability

(b) markers in terms of usability and potential impact on the marking process

(c) academics in terms of usability and potential impact on the marking process

## 1.3 Dissertation structure

The dissertation is composed of the following chapters:

**Chapter 2** provides background on design and development work conducted by students in previous years, details on related work in the form of existing systems already used at the University, and an overview of Human-Computer Interaction (HCI) user study and data analysis methods.

**Chapter 3** outlines the methodology used for the development and evaluation stages of the project.

**Chapter 4** introduces the process of gathering and prioritising the remaining requirements of the project.

**Chapter 5** describes the development activities carried out; the setup necessary for transferring the project into a new local environment; and the project's deployment on the world wide web to enable remote asynchronous user studies to take place.

**Chapter 6** explains the evaluation process as well as results obtained from user studies conducted to obtain feedback from all target user groups in terms the usability and impact of the system.

**Chapter 7** summarises the project as a whole, looks at challenges faced and gives thoughts on the future of MarkEd.

# Chapter 2

# Background

The overall project focuses on the development of MarkEd towards a fully fledged marking and feedback platform and the subsequent evaluation of the new functionalities with users.

This chapter will provide background about the systems for assignment submission and marking in UK Higher Education already in use, as well as giving a detailed description of the work done by the five students who have worked on MarkEd since 2019. The chapter concludes with a general introduction of Human-Computer Interaction (HCI) methods for gathering feedback from users and data analysis techniques. Altogether, this chapter provides the necessary background on the type of system being developed, details what work has already been done, and describes what methods might be used to evaluate whether this project has been successful.

## 2.1 Current systems used in Higher Education

### 2.1.1 Learning Management Systems

Higher education institutions now require a centralised space to maintain course documents, calendars, assessments and many more administrative tasks. A solution is purchasing a subscription for a Learning Management System (LMS). This is an online platform that supports many users (students, teaching staff, administrators) and often allows for integration with other tools such as grading platforms.

**Blackboard Learn** [56] is one example of a comprehensive LMS. It has been rated as "Best for Higher Education Institutions" in the Forbes "Best Learning Management Systems (LMS) Of 2023" list [41] and is the second most popular LMS for universities in the UK after Moodle [67][3]. Andrius Girdzius has discussed Learn in detail mentioning the system does well in handling student submissions and administrative tasks but lacks support for coding assignments [39]. In more general usability studies, Obead Alhadreti found that Blackboard Learn has "inadequate usability" due to poor design choices in the interface concerning navigation especially [9]. On the other hand, Afra Al Mansoori obtained more positive results from academics at the University of Abu Dhabi but cited

issues with intuitiveness and the need for user training before beginning to use this system [58].

### 2.1.2 Grading platforms

Grading platforms on their own provide specialised support for marking student submissions. They usually provide the ability for integration with an LMS in order for student data to be easily passed between the systems and can be thought of as an "add-on" to the LMS rather than a replacement. In this way, students can submit their work via the main LMS and instructors can mark the work within the grading platform itself, leveraging the specialised functionality and finally export the final marks back to the LMS when finished.

**Turnitin** [104] is one such platform. It is a plagiarism checker that has an extensive database of 1 billion submitted academic papers and 70 billion archived webpages [57]. It is therefore widely used for essays or longer reports; after the student submits their work, the system generates an 'Originality report' which shows text that matches sources in the Turnitin database, generating a percentage of matched text [105]. The higher the percentage, the more likely it is that the student has plagiarised (although its accuracy has been contested in the past [33]). Good features for marking, moderation and feedback include the ability to allocate a group of students to a marker. A drawback of Turnitin is that re-submissions overwrite older submissions (the only platform to do this out of the four discussed in this section) [39]. Turnitin does have the ability to easily export student data, but this is not customisable.

**Gradescope** [40] is another marking platform specialised for hand-ins, supporting AI-powered, simple rubric and auto-graded marking. The platform is best suited for marking online exams and question-based work. An advantage Gradescope has over Turnitin [104] is that it retains prior submissions allowing students to select the version they would like to have graded. Additional useful Gradescope features include allowing for coding assignments to run with an instructor-defined autograder or manual grading or a mix of the two [79], and allowing for export of student data as a csv file.

**CodeGrade** [24] was originally developed by teaching assistants at the University of Amsterdam and is specifically created for grading code assignments with written feedback or with an autograder. There are many features that make CodeGrade the go-to platform for marking, moderation and feedback of coding assignments as it has: a) command line interface access, b) custom export options using the CodeGrade Python API, and c) unlike Gradescope [40], CodeGrade allows for better integration with the main LMS and does not require grades to be transferred manually. All these features have made it a top choice grading platform at more than 50 Informatics and computer science faculties worldwide [116].

## 2.2 Previous work on the MarkEd system

MarkEd has been proposed as a system solving many problems outlined in the Motivation in section 1.1. Work was carried out by students Andrius Girdzius, Chris Sun,

Hamdani Azmi, Xisen Wang and Xiaofei Sun. The development of the project since 2019 is outlined in this section.

### 2.2.1 Stage 1: Andrius Girdzius

Stage 1 was carried out by Andrius Girdzius in the academic year 2019-20. His work centred around detailed requirements gathering with stakeholders to find out what sort of a system would be needed [39]. He conducted semi-structured interviews and organised focus groups with a total of 42 participants including students, teaching support staff, academic staff and administrative staff to create a table of main requirements shown in appendix E. He then went about creating low-fidelity wire-frames on paper to a mid-fidelity prototype supporting basic user interaction via changing screens in Figma [35]. His final prototype focused on the interface for teaching staff (senior academic, teaching assistant, marker) and included designs for the following pages:

- *Home* (fig. F.1) – The first page seen after logging in. It lists all courses that the logged in member of teaching staff is responsible for.
- *Dashboard* (fig. F.2) – The "home" page for a specific assignment. It displays information about the course teaching staff team and statistics such as the number of submissions made by students and the number of assessments left to mark for the logged in user.
- *Submissions* (fig. F.3) – Displays a table summary with all submissions made by students.
- *Modules* (fig. F.5) – Shows all available modules, which the user can select for a course. An example of a "module" could be an "Export Module", which "could be configured to export data in the format that is agreed with the ITO" [39].
- *Jobs* (fig. F.4) – Allows academics to set up a sequence of steps using one or more modules that the system would then carry out automatically. For example, automarking an assignment and then sending an email with the grade (also known as "automated workflows").
- *Setup* – The setup for the coursework in question, including marker permissions, assignment of markers and marking scheme settings.
- *Marking (data)* (fig. F.6) – A table summary of all submissions alongside marks, feedback and any tags.
- *Mark (a submission)* (fig. F.7) – This opens a selected submission to view and mark. (This is where the marking of an actual submission takes place).

Andrius' summative evaluation contained many positive comments. However, the average SUS score of 65 suggests room for improvement as participants rated the system slightly below average in terms of usability [90].

### 2.2.2 Stage 2: Chris Sun and Hamdani Azmi

Chris Sun and Hamdani bin Wan Mohd Azmi continued Andrius' work in 2021 by investigating marking strategies and feedback quality. While Chris explored techniques for producing fairer and more efficient marking [23] and Hamdani looked at strategies for producing higher quality feedback [111], both students had similar methodologies. A literature review on their topic was followed by a user study to obtain requirements. Two iterations of incorporating the requirements into Andrius's original design and evaluating these with users followed. Hamdani also designed the following pages for students:

- *Home* (fig. F.10) – The first page visible to a student after they log on. It displays a table with assignments for all courses the student is enrolled in.
- *Feedback* (fig. F.11) – Displays the student submission alongside marks and feedback received.
- *Chat* (fig. F.12) – A chat interface for direct communication with the marker of the submission.

Chris' final strategies for fair and efficient marking are summarised in table F.8. In his summative evaluation, the feature "double marking" received the highest rating in terms of impact (potential for being useful) and "small criteria-based marking" was rated as having the lowest impact. The average SUS score was 65 which was the same as the score Andrius received.

Hamdani's final list of features for fostering high-quality feedback is shown in table F.9. In his summative evaluation, Hamdani's feature "Feedback sandwich" was rated lowest in terms of impact with a score of 2.8/5 and feature "Feedback bank" was rated highest with a score of 4.5/5. Hamdani's final designs received a SUS score of 78, the highest yet moving from an "ok" usability score to "good" [11].

### 2.2.3 Stage 3: Xisen Wang and Xiaofei Sun

Using prior students' work as a basis for their contribution, in the academic year 2021-22 Xisen Wang and Xiaofei Sun started work on the code to implement a website using Python Django [28] for the back-end and Bootstrap [13] for the front-end with a MySQL [69] database. The user interface followed Andrius' design closely with the intention of adding Chris' and Hamdani's features in later. Xisen designed the overall Django template for the pages standardising the top status bar and menu sidebar [112]. He then went on to develop the general Login, Sign-up pages along with the student-specific Home, Submission and Feedback pages. On the marker side, he implemented the Home, Submissions, Marking and Mark pages (described in detail in section 2.2.1). Meanwhile Xiaofei worked on the marker Dashboard, Submissions and Setup pages [98]. She also designed the initial database schema (fig. H.12). Screenshots of Xisen's work can be found in appendix G and Xiaofei's in appendix H.

In order to evaluate the system, Xisen deployed the project temporarily on a server for public online access to the website. Overall, comments in Xisen's summative evaluation were positive with suggestions on expanding some functionality such as being able to

see files that had already been submitted. For the marker-facing pages a few markers brought up issues, such as the type of marking scheme supported being very limited and advanced filtering missing. His design was rated as having the highest impact on "Submitting assignments" with a score 4.8/5 and lowest for "time saving" with a score of 3.8/5. Xiaofei conducted think aloud studies and interviews for her summative evaluation. She received positive feedback on the simplicity of her implemented pages, but the participants highlighted the fact that guidance was missing for new users. Adding a job (automatic email reminders to markers) was the most well-received in terms of functionality, with "permission setting" having the most comments on faulty logic in the design and implementation. The average SUS score for Xisen's pages was 73 ("good") and for Xiaofei's pages it was 61 ("ok").

## 2.3 Moving forward

### 2.3.1 Web development

Web development is concerned with the development of websites and web applications that are accessible through the internet. In general, it is now commonplace to use web frameworks to alleviate repetitive or common tasks to save the developer time and prevent them from having to "re-invent" the wheel. URL mapping, database connectors and session management become much easier when deciding to use a framework. Choosing the best framework for a project comes down to a variety of aspects including the type of project and programming language [93]. Therefore, there is no one best framework to use but rather several popular ones such as Express.js [31] (for JavaScript), Ruby on Rails [89] (Ruby), Django [28] (Python) or Spring Boot [94] for Java .

A framework that is more concerned with the appearance of the user interface (UI) is also known as a front-end framework. For the appearance of the application, Cascading Style Sheets (CSS) frameworks are often used to provide the developer with standard UI elements such as forms, buttons and navbars. Popular CSS frameworks include Tailwind [100], Bootstrap [13] and Foundation [37]. Responsive elements and page layout are other important aspects that a framework such as Bootstrap can help with [43].

A database allows a web developer to store and retrieve information needed for their website in an organised, simple manner [52]. A Database Management System (DBMS) is needed to communicate and use the database effectively [85]. It facilitates reading, writing the data with security measures in place as well as concurrency control. Well-known DBMSes include PostgreSQL [78], MySQL [69], MongoDB [66]. SQLite [95] is a popular choice for smaller projects as SQLite is more light-weight but does not support simultaneous users.

### 2.3.2 Human-Computer Interaction data collection and analysis

Human-Computer Interaction is key to understanding the gap between what the designers and developers regard as a good product and what the users actually want. User study methods allow developers to prove that their product is actually useful for

the users while also finding potential areas of improvement [109]. Quantitative data collection can be used to compare results across studies while qualitative data collection allows for gaining rich insight into user opinions and behaviour when interacting with the product. With the right choice of HCI methods one can ultimately gain a deep understanding on how to improve the product to suit the user's needs.

#### Interview

The interview is one of the most fundamental survey methods for obtaining information from a participant and can be used for exploration, concept generation as well as evaluation [44]. The method involves a one-to-one personal conversation in which the interviewer (researcher) asks the interviewee (participant) a series of questions and takes note of their responses.

Depending on the goal of the researcher, interviews can be categorised in three ways: i) *structured*, ii) *semi-structured* and iii) *unstructured*. The structured interview requires that the researcher strictly adhere to the question script. The drawback of this is that no deviation or follow-up questions are allowed, but timekeeping and data analysis are made easier as a result [19]. In the unstructured interview, there is flexibility in how the researcher can respond, overall making the conversation more casual. It allows for rich, qualitative data collection and so lends itself well during the earlier stages of design when little is known about the topic being explored [110]. Semi-structured interviews offer a mixture of both types making sure that key topics are covered but also allowing for follow up questions to be asked.

#### Questionnaire

The questionnaire is another fundamental survey method [19], similar to the structured interview but this time in written form. Online survey tools such as SurveyLegend [99], Qualtrics [84], or Microsoft Forms [63] can be used to design and conduct this type of user study efficiently. Questions can vary in terms of selecting between different options and free-written responses. Techniques to gaining more rich data in terms of participant preferences include ranking options rather than selecting a single option and using Likert scale questions. Given a statement, a Likert scale allows participants to select an option on a scale that ranges from "strongly disagree" to "strongly agree" [51]. This is a common technique used by researchers to gauge the participant's attitude towards something via a questionnaire. Although being easy to distribute (for example using a mailing list), questionnaires can suffer from a relatively low response rate due to the impersonal nature of this study [88].

#### System Usability Scale

The System Usability Scale (SUS) [15] is a widely used method for measuring usability. It provides a quick assessment of usability using a questionnaire and can be used to evaluate a variety of products including software. The questionnaire has 10 questions made up of alternating positive and negative statements about the product such as "I found the various functions in this system were well integrated" or "I found the system unnecessarily complex". The user gives a score between 1 and 5 corresponding to

"strongly disagree" and "strongly agree" respectively. The scores are then converted to a "contribution score" between 0 and 4. Finally the overall SUS score ranging from 0 to 100 is calculated using the simple formula below:

$$\textbf{SUS score} = (\sum_{i=1}^{10} s_i) \times 2.5$$

where $s_i$ is the contribution score for question $i$.

The SUS has been criticised for being technologically outdated [16]. However, SUS has the advantage of being widely used and therefore the score can be interpreted as percentile of how it compares to other products [90]. The scale can also be used as benchmark of what makes a "ok" or "good" product in terms of usability by using the conversion to an adjective descriptor as depicted in figure 2.1 [11]. In the case of this project, as prior students working on MarkEd used this method, it is a useful benchmark for determining that the quality of work has not deteriorated.

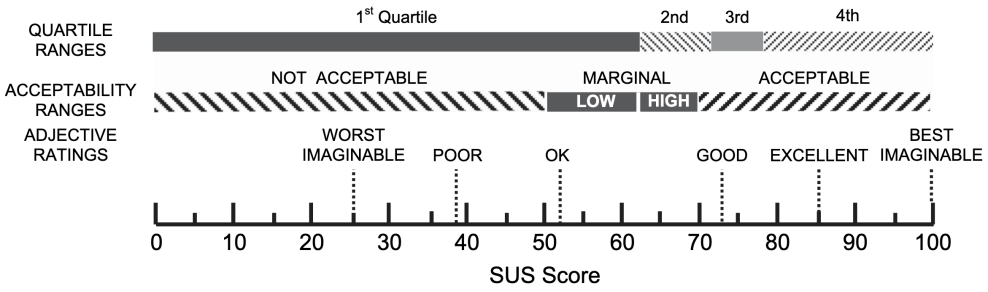



Figure 2.1: System Usability Scale with adjective descriptors [11]

#### Stakeholder walkthrough and cognitive walkthrough

Both the stakeholder walkthrough as well as the cognitive walkthrough are group usability inspection methods. They both produce qualitative results and allow for gathering evaluation data [54]. Both methods focus on setting a number of tasks (common use cases of the system) for the participants to complete individually, coming together after each task to share their thoughts. The main aim is to uncover problems hindering the user to complete set tasks rather than defending the system's design or development style.

The difference between these two methods lies in the group of participants invited. The stakeholder walkthrough involves end-users, representatives from the development team and project stakeholders (managers, investors, administrators, etc.) [80] while the cognitive walkthrough relies on HCI experts as participants. The cognitive walkthrough method is slightly more structured and is best suited for "walk-up-and-use" systems where users will want to use the system efficiently without needing any training such as ticket machines or ATMs [21]. Evaluators break down tasks into further steps and at each of these answer 4 questions related to the ease and learnability of the action sequence just carried out [19]. The stakeholder walkthrough is slightly less structured, and aims for an amicable environment for different groups of people involved with the system to give feedback and make suggestions.

### Think-aloud protocol

Think aloud is one of the most used evaluation methods [19]. It is conducted one-on-one with the researcher setting a task and the participant then giving instinctive comments on aspects of the interface that "delight, confuse, and frustrate" them [72]. It can be conducted a as a concurrent think-aloud with the participants voicing their opinions as they complete the task, or as a retrospective think-aloud, with the participant only telling the researcher about their thoughts and approach after they have completed the task in silence. Care needs to be taken when conducting a think-aloud as not every comment or suggestion from the large amount of data collected has to end up being integrated into further development of the system [73]. Additionally, researchers may find participants struggle to voice their thoughts in a steady manner throughout completing the task and so might have to provide a demonstration first.

### Data analysis methods

Depending on the type of data collected there are two common methods on how to analyse data. For *quantitative* (numeric) data, when looking at one variable, univariate statistics can be computed with measures of central tendency (mean, median, mode) and dispersion (range, quartiles and standard deviation). To examine the relationship between more than one variable, multivariate statistics can be generated using techniques such as multiple regression and logistic regression [60].

On the other hand, *qualitative* (text) data requires some thought about how best to summarise the data while also recognising patterns and trends [48]. For this purpose, thematic analysis can be applied. This technique involves the researcher reading through the text data gathered (could be transcripts of recordings or written responses by participants) multiple times. On the first read-though, the researcher "codes" the data. Codes themselves are categories that the researcher thinks a particular excerpt of text fits into. Coding can either be *deductive* where the researcher creates the codes before the read-through or *inductive* where codes are generated as the texts are being read [14]. Finally, the codes and excerpts are reviewed and the researcher starts to group them into "themes". These are often more general, and less specific to the text in question than codes [47]. In general, it is recommended to do multiple rounds of coding and reviewing themes to become better familiarised with the data, allowing for a higher quality analysis.

# Chapter 3

# Methodology

A mixed methodology was used for this project, consisting of:

1. identification and prioritisation of remaining requirements (relevant to RQ1.a, RQ1.b and RQ2)
2. development work on the MarkEd code base to implement high-priority requirements (relevant to RQ3.a, RQ3.b)
3. empirical assessment of new functionalities via user studies (relevant to RQ4.a, RQ4.b and RQ4.c)
4. summary of the project including recommendations on future work in order to launch MarkEd as a fully fledged marking and feedback tool.

The remaining requirements were identified through careful examination of all the work done by the five students who had worked on MarkEd in the past. Andrius' original requirements (appendix E) as well as all features designed by Chris and Hamdani were taken into account and then checked against Xisen and Xiaofei's implementation to see what was missing or still unfinished. Additionally, Xisen and Xiaofei's user studies were examined to incorporate all feedback on the existing system. All remaining 62 functional requirements were then compiled into a list which can be found in appendix D.

The requirements were then prioritised by using an adapted ICE prioritisation framework to gain an overview of which features had high, medium and low priority. Other prioritisation methods were also discussed as potential alternatives, with the ICE prioritisation method chosen for being relatively rigorous while also quick to apply. This is discussed in more detail in chapter 4.

Next, additional steps needed to be taken into account before the development of new features. The setup of the project on the local machine is made harder when porting from one operating system to another. With no setup instructions, the correct libraries had to be installed and problems with the MySQL [69] API drivers for Django [28] sorted out. Additionally, no database files from the previous students were available, and so synthetic data had to be created in order to fully understand the range of functionality

that the system already had.

It was decided to start development work on the student pages as this offered a good split from the other student working on MarkEd who was working on the marker pages. It also allowed me to work on implementing requests made by the largest target user group, namely students. The lower complexity in the back-end provided me with more opportunities to improve my web development skills while also allowing for more creative freedom on the front-end as well.

The system was evaluated by means of an online Questionnaire (described as an approach in section 2.3.2). This type of study was chosen mainly to obtain as high participant numbers as possible. As students make up a far larger user group than academics or markers, the sample size had to be greater to be significant. A questionnaire typically takes less time than other user studies such as interviews or think aloud studies, with the aim being 10-15 minutes to complete. While interviews do gather more in-depth data, their results are generally not comparable across studies unless the interview is structured in some way [19]. Given that a structured interview is an in-person questionnaire and that anecdotally students (especially) might find interviews daunting, the questionnaire was chosen for the evaluation with students.

After implementing all high and medium remaining requirements for the student pages, work started on the interface for teaching staff, referred to as the "marker pages". Again, the highest priority requirements were implemented, included highly requested features such as "shuffle", "custom tags", the "timer" and "bulk adjust marks". Implementation often followed the original designs from either Andrius, Chris or Hamdani in the first instance. Where there was not enough information available, development that affected the user interface were was guided by adhering to usability heuristics [74].

To allow for remote asynchronous user testing, the system needed to be hosted publicly on the world wide web for easy access. Security trade-offs had to be considered when deploying the server as barriers to entry such as restrictive firewalls on the server, though safe, could hinder participation in user studies. On the other hand, care needed to be taken when quickly deploying the project to the server, as it took some time to resolve all security checks and warnings.

The user study methods chosen for the final evaluation were the stakeholder walkthrough for the group study with academics and a think aloud with a follow-up unstructured interview for the individual studies with markers (see section 2.3.2 for details on these data collection methods). A stakeholder walkthrough was used because a group study saved time by meeting multiple participants at once. Additionally, the exchange of ideas between participants can also generate very rich data. Originally it was intended for the study with markers to be a group study as well, but due to difficulties in gathering participants, it was decided to to conduct a think aloud session with a follow-up interview to be more flexible in organising each session around the participant's schedule. Though time-consuming overall, this methodology allowed a great deal of insightful data to be gathered that might have not been collected in a group study setting.

# Chapter 4

# Identifying and prioritising remaining requirements

This chapter aims to answer research questions RQ1.a, RQ1.b and RQ2.

The remaining requirements were identified by looking at the original requirements listed by Andrius, the features developed by Chris and Hamdani as well as the evaluation chapters for all the previous students.

The method used for prioritising these requirements was an adaptation of the Impact Confidence Ease (ICE) prioritisation framework [38]. This framework was chosen due to it being a simple yet effective way of prioritising tasks [26]. Although this framework was initially used as a way to prioritise tasks in a business setting [17], it has also been used as a method to prioritise software requirements [50]. In this case, *impact* is based on how important a requirement is to stakeholders, *confidence* is based on the prior knowledge of the developer, and *ease* is the estimated time needed to implement the requirement. The final ICE score is calculated by multiplying the scores given in each category.

A drawback of the ICE framework is that it is highly subjective [26]. Therefore the original ICE framework was adapted by using a smaller scoring scale of 1-4 (as opposed to the recommended 1-10) for scoring the individual categories. A smaller range in the scoring scale allowed for the criteria to map to a unique value thus preventing arbitrary intermediate values from being used. Additionally, the first student to work on MarkEd Andrius, used a scale of 1-4 when prioritising the original high-level requirements [39]. Therefore if the identified requirement related to one of the original MarkEd requirements devised by Andrius, this was taken into account under "Impact" as well.

Using the criteria shown in table 4.1, the category scores were derived as follows:

- **Impact**: the proponent score averaged with the original priority score given by Andrius, only if the given requirement relates to a high-level requirement devised by him. Otherwise simply the propnent score.
- **Confidence**: scored based on prior web development and HCI knowledge

- **Ease**: scored based on time estimated to implement the functionality

A sample calculation for requirement 1.1 (table 4.2) is given below:

$$\textbf{ICE score } R_{1.1} = Impact \times Confidence \times Ease = 3.5 \times 3 \times 4 = 42$$

This final ICE score then guided decisions on what requirements to implement first. The maximum ICE score was taken and divided by three to assign either a high, medium or low priority to every requirement. Table 4.2 shows the high to medium priority requirements for the student pages and table 4.3 shows the top 10 high priority requirements for the marker pages. The full tables for all the remaining requirements can be found in appendix D.

| **Impact** | **Confidence** | **Ease** | **Score** |
|---|---|---|---|
| 1 proponent | Minimal or no front-end aspects, majority back-end work required and/or integration with additional tools required. | 4 weeks or more | 1 |
| 2 proponents | Some front-end and significant back-end aspects | 3 weeks | 2 |
| 3-4 proponents | Significant front-end-related aspects and some back-end-related aspects | 2 weeks | 3 |
| 5 proponents or more | Front-end-related aspects only | 1 week or less | 4 |

Table 4.1: Criteria for scoring the ICE categories

There is no shortage of prioritisation methods and frameworks developed for software engineering or business requirements. More complex but rigorous prioritisation methods such as the "Cost–Value Approach" [53] would be more appropriate in a team setting in which resources and expertise could be collated into the framework. The simpler, but still widely used Reach, Impact, Confidence, Effort (RICE) framework [46] was considered but decided against, due it requiring more detailed data than what was available in the prior student's dissertations [62]. Examples of such missing data included the experience of participants in Xisen's study, impact scores for Chris' study and complete numbers for distinguishing between proposers and acceptors in Andrius', Chris' and Hamdani's studies. This would have made it hard to judge the difference between "reach" and "impact" values in the RICE framework, and therefore ICE was trusted to produce more consistent scores.

| # | Description | I | C | E | ICE score | Priority |
|---|---|---|---|---|---|---|
| 1.1 | Student should be able to view their submitted files. | 3.5 | 3 | 4 | 42.0 | High |
| 1.2 | Student should receive confirmation email when assignment is submitted. | 2 | 3 | 4 | 24.0 | Medium |
| 1.3 | Assignments with no submissions should be at the top. | 1 | 4 | 4 | 16.0 | Medium |
| 1.4 | 12 o'clock should not be shown as noon. | 1 | 4 | 4 | 16.0 | Medium |
| 1.5 | Instruction for the assignment to be submitted should be visible. | 1 | 4 | 4 | 16.0 | Medium |
| 1.6 | 'Back' button is needed on interior pages. | 1 | 4 | 4 | 16.0 | Medium |
| 1.7 | An explanation of sandwich structure should be provided. | 1 | 4 | 4 | 16.0 | Medium |
| 1.8 | Student notifications for marks and feedback should be added. | 2 | 2 | 3 | 12.0 | Medium |

Table 4.2: High and medium priority requirements for student pages

| # | Description | I | C | E | ICE score | Priority |
|---|---|---|---|---|---|---|
| 2.1 | The "Jobs" page function should be clearer. | 4 | 4 | 4 | 64.0 | High |
| 2.2 | Shuffle questions should be completed. | 4 | 3 | 4 | 48.0 | High |
| 2.3 | Dashboard of marking progress should be completed. | 4 | 4 | 3 | 48.0 | High |
| 2.4 | It should be possible to adjust the space taken up by a submission on the marking page. | 4 | 3 | 4 | 48.0 | High |
| 2.5 | "Days Late" column should be added to the marking data table. | 3 | 4 | 4 | 48.0 | High |
| 2.6 | Timed marking should be supported. | 4 | 3 | 3 | 36.0 | Medium |
| 2.7 | Analytics on student performance should be added. | 2 | 4 | 4 | 32.0 | Medium |
| 2.8 | The system should handle "no submission". | 2 | 4 | 4 | 32.0 | Medium |
| 2.9 | The user should be able to create custom tags. | 3.5 | 3 | 3 | 31.5 | Medium |
| 2.10 | Ability to modify marks across multiple submissions at once should be added. | 2.5 | 4 | 3 | 30.0 | Medium |

Table 4.3: Top 10 high priority requirements for marker pages

# Chapter 5

# Implementation

This chapter describes implementation work done for the requirements in tables 4.2 and 4.3. Steps taken to set up the project on a local machine and server are also described.

## 5.1 Local Django project and MySQL database setup

I inherited Xisen and Xiaofei's code-base who had started implementation in 2021. This allowed me to investigate research question RQ3.a: what additional steps need to be taken in order to transfer the project onto a new local machine. As this process was very time consuming, I started a README.md file keeping track of all the setup steps which should be useful to anyone who might continue work on MarkEd in the future.

Firstly, due to programming having started in Python 3.8.3, the newest version of Python 3.8 [81] was used (3.8.18) so as to take advantages of new security patches and the guarantee that packages already used would still be functional, thus avoiding compatibility issues [82]. Similarly, the Bootstrap [13] version was kept at 5.0.1 for the same reason of compatibility with the existing code base. A new virtual environment had to be created and due to the lack of any file specifying exact requirements, packages had to be installed based on every error encountered. Some libraries had to be installed differently such as "Pillow" [83] as the project was originally developed on a Windows machine. As soon as this was done, a standard 'requirements.txt' file was produced [7].

MySQL had to be installed and set up on my local machine. MySQL Workbench [71] was additionally installed as it provides a very easy-to-use Graphical User Interface (GUI) with features such as editing values in tables directly as you would in a spreadsheet and then converting this into SQL queries automatically. The data export and import wizard features are also especially useful as they make the movement of database files (database dumps) simpler to debug and helped me a lot down the line when it came to deploying the project as well.

The main difficulties to do with setting up the database came from having issues with Django's recommended MySQL API driver [70] and the lack of a database dump from the previous students. After browsing StackOverflow [96], the problem of connecting the local MySQL [69] server with the local Django [28] server was fixed by using the

'PyMySQL' driver instead of the recommended 'mysqlclient'. The lack of a database dump from previous students was not difficult to deal with, but it did mean a lot of time had to be invested in creating synthetic data to understand the full functionality of the webapp at this stage. This included adding users with various permissions, at least one course, some assignments, marking schemes, marks awarded and more.

## 5.2 Iteration 1: working on the student-facing pages for MarkEd

This section describes the implementation of the requirements in table 4.2.

### 5.2.1 Viewing submitted assignments

Viewing submitted assignments was the highest priority requirement as both markers and students requested this feature. The reasoning given was students often like to submit an earlier draft in advance of the deadline, or want to make small changes and re-submit [39]. Feedback gathered by Andrius also suggests students like to have the chance to double-check which version of the assignment is their latest submission and will therefore be marked. Additionally, academics supported this feature in order to prevent misunderstandings where students are sure they have submitted but the academic has not received their submission which can result in an unfortunate zero marks for the student.

Development took place on the "Submit" page. Previously there was only a submission box on the page (fig. G.8). The main idea was to use the large free space on the "Submit" page more economically by structuring the page such that the user had all the relevant information for submitting as well as viewing previous submissions in one place. Firstly, the purple sidebar on the left was replaced by the simpler horizontal navigation bar. This was done because the student interface has few separate pages and a very flat page hierarchy (fig. G.2) which meant that all pages are reachable by the main table on the student home page, eliminating the need for a classic menu. Secondly, the requirement 1.5 (table 4.2) was met by introducing the course name and assignment title on the left and moving the submission box to the right side.

While adding a separate page that lists all previously submitted files was considered, this was decided against as it would quickly become cluttered and require the user to scroll through an ever-growing list to find past uploaded files for a certain assessment. It would add complexity to the simple page hierarchy and it would also make it harder to embed a pdf viewer in a natural place or mean that users could only view titles of submissions, and would need to download the file if they wanted to view it.

Considering the points made above and to easily connect the use cases of a user wanting to re-submit a file if they are not happy with the previously submitted one, the option to check a prior submission and submit a new file on the same page was chosen (see fig. 5.1). Therefore an embedded pdf viewer was added to the page with a drop-down to select any previous version. This also reflects how standard platforms used by the School of Informatics combine the upload and view submission functions by usually

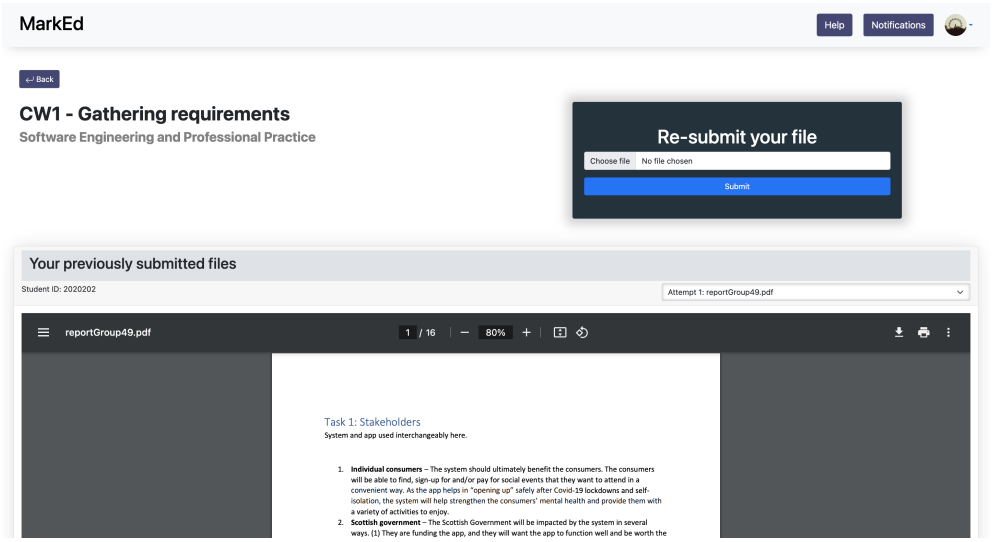



Figure 5.1: New Submit page

displaying them on the same page [97]. As such, this page has been re-structured to adhere to the Nielsen heuristics of external consistency and visibility of system status [55].

### 5.2.2 Email and in-app notifications

The second most highly requested functionality was requirement 1.2 and 1.8. Based on research by Andrius, students like to receive notifications when a submission is made to reassure them that their submission was indeed saved in the system [39]. Additionally, they mentioned that it is hard to find out when marks are released and therefore wanted to receive email notifications as well.

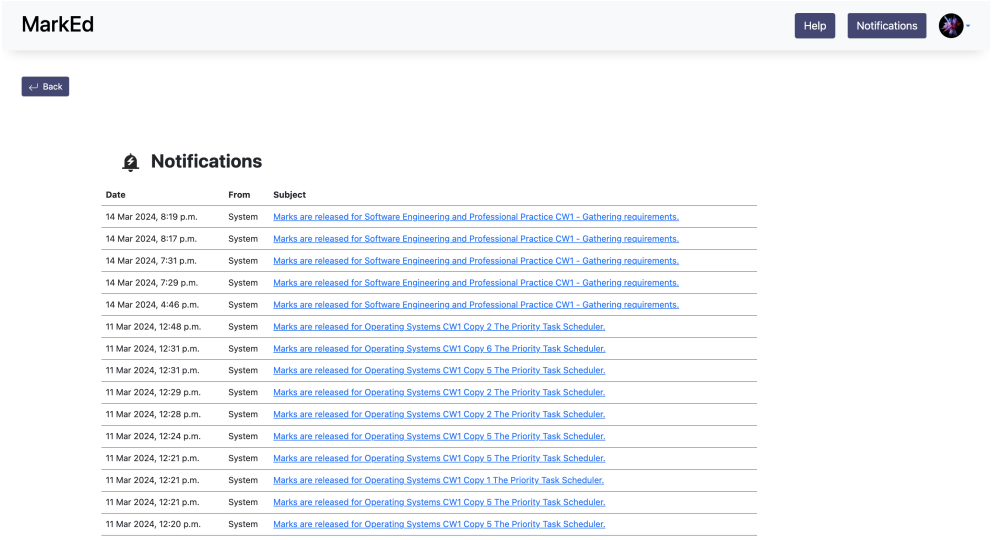



Figure 5.2: Student Notification page

The design of the Notification page seen in figure 5.2 was standardised to match the

marker notification page implemented by Xisen. Only small additions of layout and headings were changed, including the content of the table and making sure that the links in the notification indeed take the student to the submission that the notification is referencing. The last thing to take care of was to make sure a notification record was added to the databases once marking a submission was complete. Lastly, the database had to be queried every time a user navigated to the notification page to check for updates.

To implement this feature a new email address was created rather than connecting a private email addresses and risking it being compromised via the public code repository. For a new email address, email service providers often restrict users from generating an app password without seeing signs of human-like activity and so it took 5 days of using the new address until it could be connected with the Django back-end via app-authentication. After this, it was relatively straight-forward to set up the Django email settings taking care to use the most up-to-date Yahoo email Simple Mail Transfer Protocol (SMTP) server settings [101].

After the setup was complete, the code was edited such that an email is sent out directly after a submission is uploaded or after marks are released. The hardest part here was deciding where to add the email and notification code within the marker pages. It took me a while to figure out that currently results are displayed to a student as soon as a marker selects “Complete” when they have finished marking one submission. This means that students receive their marks directly after their own submission has been marked rather than when all the students on the course have had their submissions marked.

### 5.2.3 General usability improvements

The remaining requirements (1.3, 1.4, 1.6 and 1.7) from table 4.2 were also implemented and screenshots relating to all implementation in this subsection can be found in appendix E. Discussion is kept brief due to page restrictions. Following user feedback a “back” button was added to make it easier for the user to navigate back to the previous page they had been viewing. It has been added to all the interior pages (all pages except for the Home page). A Bootstrap popover [77] was added to clarify the “feedback sandwich” structure to inquisitive students while avoiding clutter on the page.

## 5.3 Iteration 2: working on the teaching staff-facing pages for MarkEd

This section describes the implementation of the requirements in table 4.3.

### 5.3.1 Dashboard and student analytics

The dashboard is the first visible page when an instructor selects a specific assignment on their home page. It is vital that this page is informative and provides a good overview of the assignment marking progress at a quick glance. It already had the basic elements

implemented by Xiaofei (fig. H.1) based on designs by Andrius (fig. F.2). Based on comments made by users, the page was re-designed to utilise the free space more effectively (see fig. 5.3).

Firstly, the marking page method had to be refactored to make sure all of the statistics under "Submissions" (total submissions, completed marking, and submissions flagged for help) were computed correctly. For each statistic generated, more advanced filters and Django's in-built aggregation methods [5] were used instead of simple nested loops iterating over an entire set of objects retrieved from the database to increase code readability. The dashboard HTML template was also corrected such that the correct profile picture appeared for every instructor on the course as opposed to a copy of the profile picture of the currently logged in user for everyone (which was previously the case, see fig. H.1). This issue was easy to fix as it required changing a line of code that took the current user's profile picture each time to instead use the profile picture of the 'User' objects passed to the template view from the back-end.

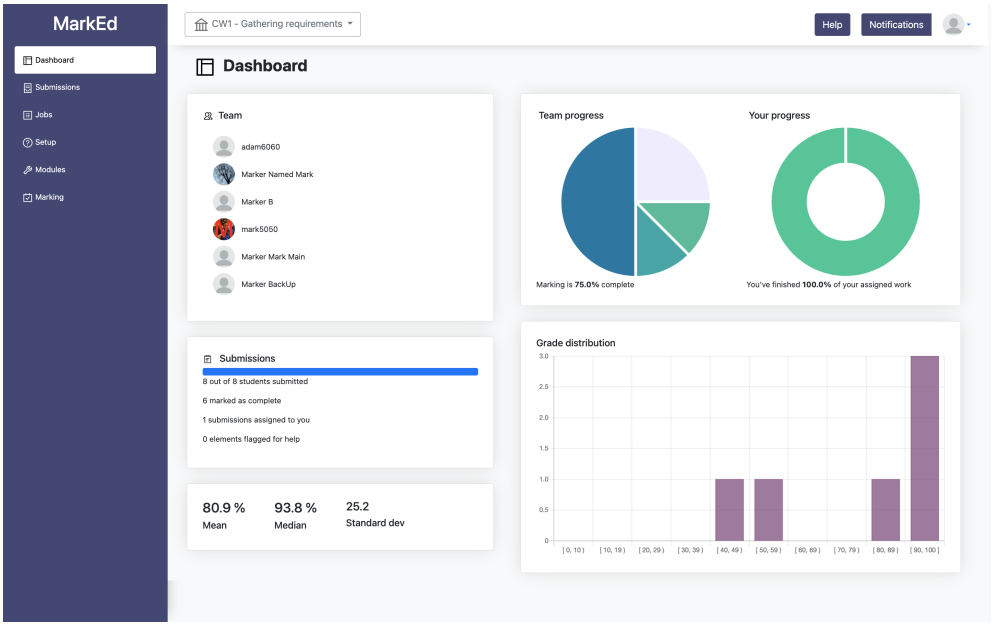



Figure 5.3: Re-designed dashboard with statistics, grade distribution chart and marker progress breakdown

The 'Team progress' chart was converted from a doughnut chart into a pie chart to differentiate it more easily from the 'your progress' chart. It now includes counts of submissions marked for every instructor. Previously it simply displayed the cumulative number of submissions marked across the team.

Two common elements in other marking platforms were added including a progress bar to view how many of the students on the course have submitted their work as is done in Blackboard Learn [56] as well as a grade distribution as is used in GradeScope [40]. Statistics and a grade distribution were added as these were mentioned as being a useful feature according to academics in Andius' dissertation [39]. Having the grade distribution as a histogram would also improve the speed at which information can be taken in [34].

To create the visualisations, the open-source Javascript library Chart.js [22] was chosen due to its simple design and customisation options. Additionally, its well-maintained documentation meant it was quite easy to look up how to create charts with an uncertain number of data points. To dynamically create the charts, the calculated statistics were saved into JavaScript variables to be then used as data points and labels. A tooltip displays the marker name and marked submission number when the user hovers over a section in the pie chart to make sure that information is not only conveyed by colour, adhering to the W3C guidelines [107]. The colour palette itself was created using the ColorSpace generator [25] based on the MarkEd brand purple used heavily in Xisen's design which can be seen in the sidebar on the left in figure 5.3.

### 5.3.2 Shuffle by question

The "shuffle by question" functionality (Requirement 2.2) was first designed by Chris as a way to reduce bias that can arise when always seeing submissions in the same order [23]. The "Shuffle" button was placed next to the the "Previous" and "Next" buttons at the top of the 'Mark' page (fig. 5.4), adhering to the gestalt principle of proximity (grouping related elements) [108].

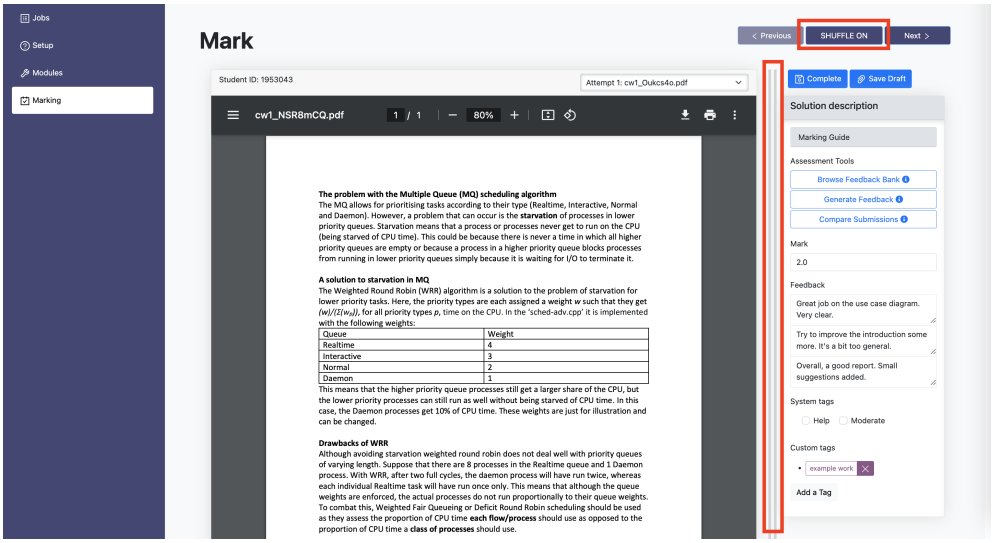



Figure 5.4: Shuffle button at the top, panel width slider identified by double vertical lines

A decision had to be taken on which submissions to shuffle. Assuming that the original order of submissions is based on the student ID, the two options were:

1. shuffling all submissions, regardless of the submission currently being viewed
2. shuffling all remaining submissions in the original list only (all submissions by students with a higher student ID number).

At first I instinctively thought of "shuffling" as the the first option, which is widely used in streaming services to create random-order playlists [113]. However, the user might already have done some marking and so they would not want to see submissions they had already marked when they click "Next". Therefore the second option was chosen.

One difficulty that arose was getting the "shuffle" mode to stay on until the user turns it off. I ended up finding out that this was because I was not saving the shuffled list of submission IDs anywhere. To solve this, the shuffled ordering of assignments is saved server-side when the "Shuffle" button is clicked. Once shuffle is turned on, I used `shuffle=True` in the Django template so that the "Previous" and "Next" buttons' href (link destination) attribute also include `shuffle=True`. This triggers the back-end to use the saved shuffle order to get the correct submission. When the "Shuffle" button is clicked again to turn shuffle off, the order of assignments returns back to the original order.

### 5.3.3 Custom Tags

Custom tags are a feature designed by Hamdani that aim to create better quality feedback as a result collaboration between markers [111]. The idea being markers can leave small notes ("tags") so that other markers or senior academics can help out if necessary. For instance, a tag can be created seeking clarification on how to apply the marking scheme for a particular submission. Users also wanted to be able to leave notes for themselves or to categorise submissions in some way [112]. A "TagCustom" table was added to store these in the database.

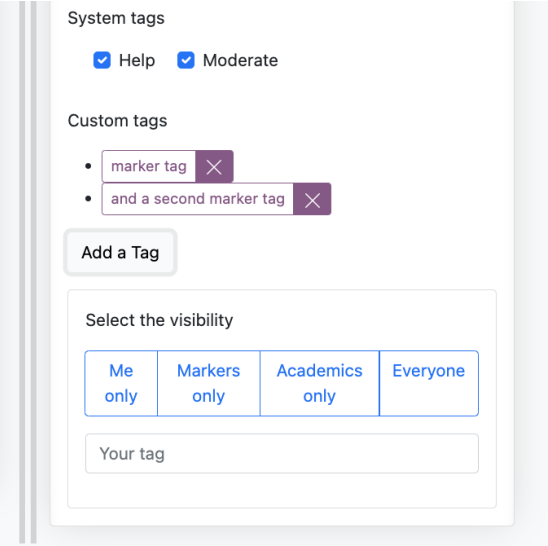



Figure 5.5: Custom tags drop-down with visibility options

Figure 5.5 shows the addition to the marking panel such that custom tags can now be created and viewed. The "Add a Tag" button expands the input section. Initially, this section is collapsed to save space and to not detract from marking. Additionally, the user can determine which visibility setting to add for the tag in question. A personal tag would be set to be visible to "Me only", whereas a tag mentioning waiting for clarification on the mark scheme could be set to be visible to "Everyone". Support for adding multiple tags at once by separating them in the text input box with a semi-colon symbol ';' was added based on a suggestion from the supervisor of this project. Tags were added as specially formatted buttons, such that an 'X' to the side of the said tag indicated the possibility to delete the tag in question. The Bootstrap "Button group" [20] was chosen rather than classic checkboxes for an overall more aesthetic design and to dstinguish it from the "System Tags" visible at the top in figure 5.5.

There was an issue of not being able to create nested HTML forms. A day after I had finished work on the "custom tags" I realised that the "Complete" button was not working due to nested forms behaving unpredictably and overall being described as bad style [103]. After finding this out, it was solved by deleting the inner form making the

custom tags part of the overall form for the submission marking. The tag fields had to be made optional rather than required, to ensure users could add marks and feedback while optionally creating a tag.

### 5.3.4 Timer

The timer concept (requirement 2.6) was first developed by Chris as a strategy for fostering more efficient marking [23]. The idea being that having an in-built timer within the website itself along with some statistics on the marker's progress would help with personal time management.

At first, this feature seemed easy to implement. An initial search on a web browser quickly displays some small timer projects made purely in CSS and JavaScript [45]. They acted as good inspiration for the initial development of the ribbon across the top of a page (shown in figure 5.6), as Chris had originally designed.

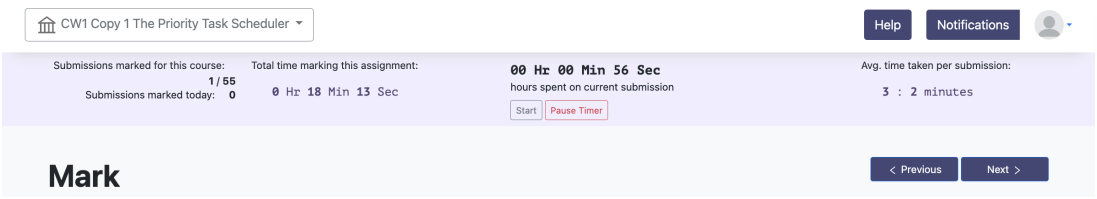



Figure 5.6: Timer ribbon with timer and marker statistics

However, simplicity was quickly lost when taking into account the website back-end and additional database tables needed in order to store time data. The user was not supposed to only have the possibility of timing themselves, but also view statistics on their progress and average time taken to mark a submission. This required storing the time data somewhere. This caused quite a lot of problems due to not fully understanding the interaction between JavaScript [49], Django [28] HTML templates and the Django Python back-end. Due to this, the implementation revolved around using the "Start" and "Pause Timer" buttons to trigger the back-end to save the time as this was the only familiar way of the user triggering a process on the server at the time. So, when the user clicked "Start", a POST request was made to the server, then the timestamp was saved in the database table "TimerRecord" along with an indicator of which button had triggered this (in this case "Start").

Additionally, to the "Start" and "Pause Timer" buttons, there were edge cases to be handled with regards to timer usage. Firstly, the timer had to start when the user first starts marking, as included in the original designs and following the idea that the user does not consciously have to remember that they need to start timing themselves. Therefore a JavaSript `window.onload()` function was added to handle this case [115]. The timing data also had to be saved when the user left the page as well. Otherwise, a marker could have been marking for half an hour, used the timer to keep track of time but this then would not show up in their statistics. It would not be a very user friendly design to expect the user to press "Pause Timer" every time before leaving the page. To add this functionality, the JavaScript Beacon API was used to initialise asynchronous requests [12]. This was achieved by creating an event listener for `visibilitychange`

and meant that as soon as the user had navigated elsewhere this triggered a request to the server to save the timestamp.

Finally, the statistics had to generated. Figure 5.6 displays the ribbon including 4 segments from left to right: **1)** submissions marked for this course and submissions marked today, **2)** total time taken marking for this assignment, **3)** the timer itself, **4)** average time taken to mark one submission for this assignment.

All of these statistics are personalised to the user. Relying on time-stamping user interactions with the buttons, meant that the duration had to be calculated as well. This was done during the handling of a GET request for the "mark" page and stored in the "TimeDuration" table. This aimed to make sure that at every re-load of this page, the marker time and progress statistics were up-to-date.

The main reason for these implementation choices was lack of in-depth knowledge of how to store state appropriately client-side. Unknown at the time, this approach later caused the annoying property that the page had to reload every time the timer was manipulated. This was by far the hardest feature to implement and as such has the most room for improvement in terms of functionality. In hindsight, making use of session cookies to store the user's interaction with the timer and finding a less disruptive way of sending the "time data" back such as only on when the marking "complete" button is pressed, or when the user navigates away from the page would be a better way to go [106]. This would also eliminate the need for a "TimeRecord" database table and instead having the "TimeDuration" only.

### 5.3.5 Bulk adjust marks

Bulk adjusting marks is a requirement found by Andrius when developing the original concept for MarkEd. This feature follows the original designs quite faithfully. Firstly the "Bulk Assign" button was added to the "marking" page (fig. E.9). Selecting this button opens a Bootstrap modal [65] which can be seen in figure 5.7. This allows for selecting the students to adjust the marks for and then choosing methods for adjusting the mark. These are: **1)** scale by factor, **2)** enter raw mark, and **3)** enter percentage (University Common Markscheme grade [32]). These options were chosen according to findings in Andrius' dissertation on how marks are adjusted as part of the moderation process at the School of Informatics and to give academics flexibility in how they wanted to use this feature. If a value is entered causing a mark to be less than zero or more than the maximum for this assignment, the mark is changed to zero or the maximum respectively. Changes are saved using the "Save changes" button which also exits the modal and brings the user back to the "marking" page (fig. G.5).

### 5.3.6 Further additions

Requirements 2.1, 2.4, 2.5 and 2.8 (table 4.3) were smaller scale but still highly requested requirements. Descriptions of these added features is kept brief due to page restrictions in this dissertation. The jobs page was re-deigned to be "clearer" as requested by 5 users. See figure E.8 for the re-design. Adjustable panels sizing (fig. 5.4) was the most requested feature with 8 proponents across the years and aimed to provide

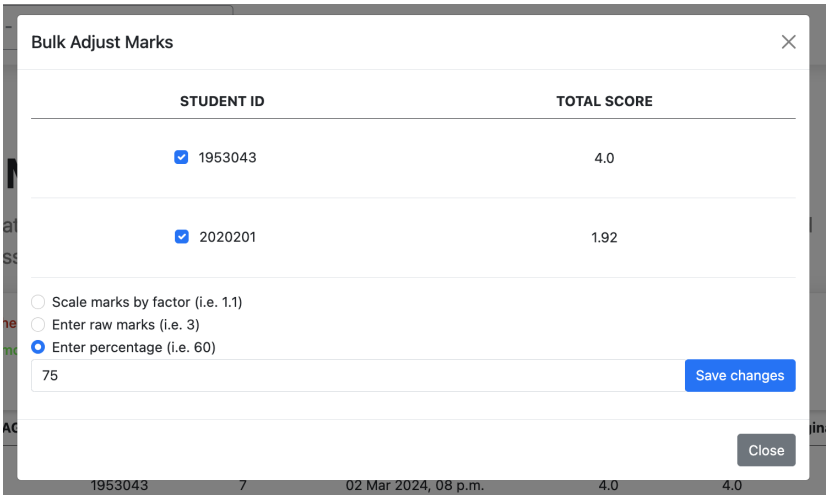



Figure 5.7: Modal for bulk adjust marks

a more usable marking interface through customisation opportunities [91]. “Days late” and “No Submission” (NS) were added to the table on the Submissions page (fig. E.12).

## 5.4 Deploying the project for user testing

To allow for remote asynchronous user testing, the system needed to be hosted publicly on the world wide web for easy access. This section therefore investigates research question RQ3.b: what additional steps need to be taken in order to host the system online for user testing.

After finishing my first iteration of implementation, my supervisor put my colleague and myself in contact with School of Informatics technicians. The technicians gave us a Linux virtual machine on a University server with one MySQL MariaDB [59] database. For security reasons, they were apprehensive about the server being open on the world wide web. However, at the time it was my understanding that the server would be accessible from the university network Eduroam [30] and the University’s Virtual Private Network (VPN) FortiClient [36].

With a new machine, come new setup problems. There was trouble at first connecting to the database. After an email exchange with the technicians I found out that the Informatics server runs on a socket, whereas my local installation of MySQL runs on a port. I was not even aware that this was a possible cause for the error I had been receiving when trying to connect to the database.

At this stage I also encountered a problem that when importing a dump file, it restores the “original” database [87]. My colleague and I were effectively overwriting each other’s work when uploading our dump files. Manually changing primary keys (PKs) for every table, was not a good idea as there was too much data, and in order manage this, we would have had to coordinate our PKs and do a lot of work in our local databases. It would also disadvantage the person who had to take the “higher” PKs, as one person would keep their PKs in the range of 1-1000 as they already were, while the other would

have to change everything to start at 1001.

Although I later received my own database from the technicians, and did not have to worry about overwriting my colleague's data, the fact that the server was only accessible via the Informatics VPN OpenVPN [75] meant I decided to look for other options. This was because having the server be inaccessible from even the Univisity network or the more widely used Univeristy VPN FortiClient [36] meant that even my supervisor could not access the live website. I could not risk students not having the Informatics VPN installed, as this would mean I would have to cancel my user study.

After I managed to obtain a private Microsoft Azure [10] Linux server, I had to start from from the beginning, installing MySQL and following all the installation steps from section 5.1. With no firewall around the server, I could also connect to the database via MySQL Workbench [71], simplifying any quick modifications I had to make to the database.

After 4 weeks, the website was finally reachable via the server's IP address (fig. E.15).

# Chapter 6

# Evaluation

This chapter contains the descriptions of all user studies conducted as well as a discussion of their results. Although conducted at different times, both the formative and summative evaluations are placed in this chapter to allow for comparisons between results of different user groups and to provide one comprehensive evaluation of the development work undertaken in this project. All of the conducted studies aim to answer the final three research questions RQ4.a, RQ4.b and RQ4.c, in order to see how the target end-users react and gain valuable information on how the changes impacted their experience.

## 6.1 Evaluation with students

The aim of this study was to investigate RQ4.a: how are the new changes perceived by students in terms of usability. This was done by means of a questionnaire that asked about students' experiences in submitting assignments and viewing feedback.

### 6.1.1 Data collection methods

Data was collected by means of an online questionnaire (this approach is described in section 2.3.2). The reasoning behind choosing this type of study is given in chapter 3. To conduct the study itself, the platform Microsoft Forms [63] was used due to the ease of setup using the university's Microsoft account and the platform's adherence to data storage compliance. The name and email of respondents was not collected to minimise the amount of personal data collected.

The SUS [15] questions were incorporated in every questionnaire for every user group. It was important to include these because they offer quantitative scores that can be compared across different systems and over time [16].

### 6.1.2 Participants

Participants were sourced by a variety of methods. Firstly, my supervisor reached out to the student body at the School of Informatics via various student mailing lists for the

different study years (including MSc taught students). Additionally, I participated in user studies myself, so that the other students could participate in mine as they were also having trouble sourcing participants. Lastly, acquaintances reached out in their circles to interest a few more people as well. The list of participants and their year group can be found in appendix C.

### 6.1.3 Materials and procedure

Following the University's ethics and data protection guidelines, participants were presented with the relevant Participant Information Sheet (appendix B) and Consent Form (appendix A) before starting the questionnaire. They were also given a link in order to access the MarkEd website on their personal laptop or computer.

The questionnaire was structured into three parts:

**Part 1**: Questions centring around the completion of tasks following a typical cycle of 1) submitting an assignment, 2) being notified of marks release, and 3) viewing marks. The tasks to be completed were uploading a pdf to a submission box, viewing a previously made submission, viewing feedback on a marked assignment and viewing notifications.

After every task, a series of questions were asked. These included whether the participant had successfully completed the task, in order to measure the completion rate of tasks as an important aspect of usability [1]. Questions about the content, presentation and layout were also asked each time to assess how well changes helped users to complete their tasks on the website [2].

**Part 2**: General suggestions for the system and measuring preferences in the interface. This included improvements that could be made to the overall style and visual presentation of the interface, as well as the types of notifications users would like to receive if the system were to be trialled at the School of Informatics in the future.

**Part 3**: 10 SUS questions as specified by the original paper [15] with one terminology change (”people” changed to the target group ”students” in question 7).

Suggestions for potential improvements were also collected. Test questions were added in order to determine whether to disqualify participants for not reading the questions and answering randomly [68]. The full list of questions asked can be found in appendix C.

### 6.1.4 Data analysis methods

Task completion rate was calculated as the percentage of the total number of participants who successfully completed a task. All quantitative data collected was analysed using descriptive statistics and measures of central tendency (see section 2.3.2) to compare which pages were most positively and negatively received, as well as the spread of opinions. Additionally, the long text responses were read and analysed for multiple mentions of the same aspect to see what the users as a whole found positive or negative. For the exploratory questions in part 2 of the questionnaire, the percentages of the user

groups who preferred one option over another were compared to give an indication of the potential focus of future work on MarkEd. SUS scores were calculated using the standard formula described in section 2.3.2.

### 6.1.5 Results

#### Task completion rate

The completion rate for tasks was high, with an average of 92.8% as can be seen from table 6.1. The fact that task 3 and task 4 were completed 100% successfully shows that the implementation the feedback and notification pages worked as expected. The highest amount of unsuccessful task completion was observed in task 1: submit a file to the submission box. The reported error was "Server Error 500". After inspection, this uncovered a bug where clicking "submit" without selecting a file resulted in the page refreshing incorrectly. On the other hand, participants who successfully submitted the file reported that they noticed the confirmation message after submission, signalling that this type of system feedback message can be used throughout the system in the future.

| Task 1 | Task 2 | Task 3 | Task 4 | Average |
|---|---|---|---|---|
| 81% | 90% | 100% | 100% | 92.8% |

Table 6.1: Student task completion rate

Task 2 (viewing a previously submitted file) was particularly interesting, as this feature was completely new and had not been tested with users previously. Originally, it was thought that navigation might be an issue. However, 90% of participants had no trouble navigating to the "submit" page to view their previously submitted files. The remaining 2 students commented on an error they received. This was an error with hosting the website rather than the design or implementation itself, as there were difficulties with refreshing static files which probably caused this error for the participants. This would be a concern when planning more advanced acceptance testing, but it is not strictly speaking relevant to this study as it is purely a deployment issue until all main functionality has been completed.

#### Page-specific user experience

Results concerning the content, presentation and interaction aspects of the student pages were positive (full results can be found in appendix C). The "submit" page was rated highest with 89% of participants giving a rating of "good" or "very good" on content and presentation and 100% on interaction. However, S16 did mention a problem with responsiveness to screen size:

*"The layout scales reasonably well with smaller screen sizes compared to the main screen, but for some reason the main navigation disappears when screen width is less than ∼1000px, this is probably unintentional behaviour." — S16*

Around 18% of participants gave the content on the "feedback" page a rating of "poor" or "fair" and 23% gave the presentation a rating of "fair". The main negative point seemed to be the feedback itself. This might have been an oversight when designing

the tasks, as the sample feedback displayed was not realistic. This meant many free response suggestions centred around the feedback rather than the website page. 5/21 students commented on the text size and formatting of the "feedback" page as being bothersome.

#### Other findings

Participants were receptive to the suggestion of expanding notifications, contrary to previous studies by Andrius [39]. Unsurprisingly, students stated that "Marks have been released" was the most important notification type. 6/21 participants had additional suggestions, such as reminders before approaching deadlines or related to changes in deadlines. Findings suggested that the notification types are worth expanding but also allowing customisation would be a key aspect in future work.

## 6.2 Evaluation with teaching staff

The aim of this study was to investigate research questions RQ4.b and RQ4.c with academics and markers. The studies were constructed as a series of typical tasks found in marking.

### 6.2.1 Data collection methods

The final evaluation used a group study with mainly academics in the form of a stakeholder walkthrough (see section 2.3.2) and an individual study with markers comprised of a concurrent think aloud (see section 2.3.2) with a follow-up unstructured interview (see section 2.3.2). The reasoning behind choosing these types of study is given in chapter 3. All studies were conducted synchronously online via the platform Microsoft Teams [63]. This also allowed for automatic recording and transcription of the meetings. Due to a mistake in turning off the automatic transcription feature, the stakeholder walkthrough had to be transcribed separately.

Data was also collected in the form of an online questionnaire using Microsoft Forms [63]. Questions about impact on the marking process were nearly identical to those devised by prior students, in order to ensure that results could be compared across years. As mentioned previously, the SUS [15] questions were incorporated in every questionnaire for every user group. They were included because they offer quantitative scores that can be compared across different systems and over time [16].

### 6.2.2 Participants

The project's supervisor was very helpful with the organisation of the group study. In this way, the stakeholder walkthrough with five academics and two markers could be conducted during a teaching staff meeting. This was greatly appreciated as it saved a lot of time trying to gather participants and agreeing upon a time.

At the same time, there was a lot of trouble sourcing participants for the study with markers. Gathering participants via the general mailing list for all student-markers

was unsuccessful as only one marker signed up. Therefore, various distant contacts who were markers themselves or might know markers had to be contacted by email directly. After several days three markers in total had been found, meaning that in total 5 academics and 5 markers took part in the user studies.

The list of participants and their experience marking (and acting as course organiser if applicable) can be found in appendix C.

### 6.2.3 Procedure and materials

Following the University's ethics and data protection guidelines, participants were presented with the relevant Participant Information Sheet (appendix B) and Consent Form (appendix A) before starting the user study.

#### Stakeholder walkthrough

The meeting started with a short introduction on MarkEd and the plan for the study. Next, the participants were given login credentials in order to log into MarkEd and open their own assignment. This part took around 15 minutes to complete.

After the users had all found their way to the "mark" page, the users completed 4 tasks. For every task, the stakeholder walkthrough was broken down into the following steps:

1. Researcher presents the task on the PowerPoint [64] slide.
2. Participants have 5 minutes to complete the task. Tasks included using custom tags, the timer, the dashboard and bulk adjusting marks.
3. As participants work through the tasks they voice their first impressions.
4. The researcher then opens up a discussion with some prepared questions. The supervisor for this project also helps with moderation. This step takes around 5 minutes.

The times given are rough estimates, as some tasks generated less or more discussion than anticipated, in which case the group discussion could last up to 10 minutes. Details on tasks and questions can be found in appendix C.

As the entire meeting lasted 1 hour in total, it was not possible to ask participants to complete the questionnaire during the study itself. In the end, 2/5 academics completed the questionnaire in their own time.

#### Think aloud and unstructured interview

The meeting started with a short introduction to MarkEd. Next, the Think Aloud Protocol was described to the participant. This was done to clarify that the participant needed to constantly talk without any response from the researcher, as some people find this very unnatural [73].

Every meeting followed this plan:

1. **Think aloud** (20 minutes): The participant completes a set of tasks while constantly talking about what they are doing and thinking. The researcher's only input is to remind the participant to talk.
2. **Questionnaire** (10 minutes): The questionnaire was built in as part of the study to avoid people forgetting or not having time later.
3. **Unstructured interview** (10 minutes): The interview served as an opportunity for the researcher to ask follow-up questions on something they had noted as interesting during the think aloud. A set of general questions was prepared as well, only some of which were asked due to time constraints.

The questionnaires for the academics and markers contained identical questions about the impact both the new features as well as the system as a whole could have on marking at the School of Informatics. All of these questions were answered using a Likert scale. Questions on the impact of new features added was asked to identify whether any unexpected patterns emerged, in which users found a specific feature more useful in a different area than originally intended. For example, a feature being rated highly for "saving time" even though it was supposed to help produce high-quality feedback. The questionnaire concluded with the standard SUS [15] questions.

### 6.2.4 Data analysis

The user studies with academics and markers mainly generated qualitative data. Therefore data was analysed using thematic analysis using NVivo 13 [4]. For the questionnaire, the Likert scale responses were converted to scores 1-5 and these were then analysed using descriptive statistics and measures of central tendency.

For the thematic analysis, all the transcripts from the almost three hours of recordings of the stakeholder walkthrough and think aloud sessions with subsequent interviews were read and analysed to gather initial codes for the topics that users discussed. A mixture of deductive and inductive coding was performed [14]. The deductive codes were pre-determined, based on the features with which the users had interacted. More specific aspects of these features that were deemed significant were added as sub-codes. Comments made that did not fit into specific features but were connected to usability in general were added using inductive coding. Suggestions were also coded. The initial codes concerning usability can be seen in figure C.1, with the deductive codes shown in blue on the left and inductive codes shown in orange on the right. The size of the boxes shows the relative frequency with which these codes occurred in the transcripts.

The codes were then translated into themes by reviewing all the snippets of text falling under a certain code. The themes discovered can be seen in figure 6.1. For more information on the data analysis methods used, please see section 2.3.2.

### 6.2.5 Results: Thematic analysis

Overall, a great deal of feedback was collected, both positive and negative, with the group study offering more general feedback and the think aloud sessions offering insight into smaller details in the interface that made a difference while marking.

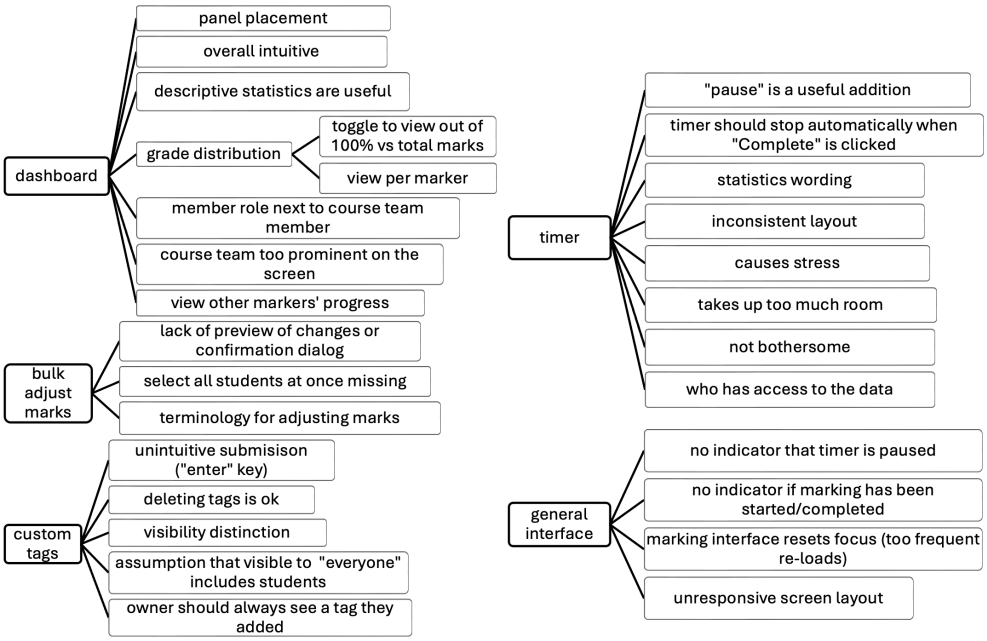



Figure 6.1: Themes identified from the user studies with teaching staff

All markers worked out how to add **custom tags** by pressing "Enter" on the keyboard relatively quickly, but 3/5 markers and all academics wanted an "Add" or "Submit" button. Deleting tags did not pose a problem for the 10 participants who interacted with this feature.

It was mentioned that the visibility ("private", "academics only", etc) of a tag was not clear. A5 suggested distinguishing the visibility level of a tag by colour, A2 mentioned grouping and M4 mentioned using icons.

Regarding the functionality of the **timer** feature, A3 mentioned they expected the timer to stop automatically when "Complete" is clicked on the Mark page (top right in figure 5.3.2):

*"when I've completed it [marking the assignment], I think that could be where the timer could finish because it'll tell us how long I actually spent marking as opposed to just looking and browsing through the documents and just looking at the assignment." — A3*

A critical bug was also uncovered:

*"I later realised 'pause timer' is lying to me: it means 'pause the display, the timer keeps running'" — A2*

A reason for this could be last-minute changes to the timer undoing the correct functioning of the "Pause" button.

For the statistics generated, A1 wanted a more standardised visual layout in the timer ribbon, with text at the top and the digits at the bottom.

The timer resulted in a large general discussion on time-keeping during marking in the group study. A4 mentioned several times that they would prefer to not have a timer at

all. A1 concurred. This led to a suggestion to enable the timer to be hidden. This would also free up space on the screen for laptop users:

*"Just the ability to shrink that ribbon so it's not in your face the whole time would be nice." — A3*

On the other hand, a less experienced marker M3 was very receptive to the timer:

*"Maybe if I was going through like a massive amount of like exam submissions it would be useful, but I think I would be much more likely to use the timer in the interface for sure. [...] I think it would be really useful." — M3*

A3 was also interested in whether the timer could be used for marker analytics. Originally, this feature was just designed for personal time management but A3 mentioned it would be useful to have access to the statistics for course management and budget allocation. This sparked a debate on privacy and stress caused by having your marking times visible to others. The consensus reached was that anonymised data could be used.

Overall, 3/5 academics and 2/5 markers mentioned they would use the timer, with 2/5 academics stating they would not. The remaining 3/5 markers made no comment.

The **dashboard** was the least controversial. As M1 summarised it:

*"It's intuitive. It's not too busy, which is something that I've seen in marking tools like this before." — M1*

There were some suggestions on how to improve the dashboard. The most common complaint was that the grade distribution (see figure 5.3 bottom right) histogram x-axis corresponded to the University Common Marking Scheme percentage grade [32] rather than the raw mark awarded. 3/10 participants agreed that this was bothersome and A5 suggested a change should be made to enable users to toggle between two views.

A more complex solution that might satisfy most users was given by A2:

*"You'll never find one layout that suits everyone, so [you] could make each component be rearrange-able by the user (and the arrangement saved)." — A2*

Generally, the academics completed the **bulk adjust marks** successfully. The largest source of confusion was in the wording of how the academic wanted to adjust the mark (such as scaling by a factor, or giving a raw mark).

3/5 participants commented that having a preview of the adjusted marks or having a confirmation dialog before the marks are changed permanently is needed. A1 summarised this:

*"When you do that adjustment, maybe it's good to have a kind of warning window, some small window, asking you again or even showing you what the marks become before you click 'Save Changes'." — A1*

Additionally, A5 suggested a "Select all" (students) option, to which A4 agreed.

The most common **general issue** users faced was constant reloads of the "mark" page. As M4 said:

*"It's refreshed the page and closed this. I don't love that. It just resets my focus [to] where I was already."*

The most likely cause for this is the prevalent use of synchronous requests to the server for all the buttons on the "mark" page. As more functionality is added, this degrades performance from the user's perspective as every interaction triggers a request to the server which takes up valuable time.

3/3 markers successfully located and used the **shuffle** feature. Since users did not face any challenges when using the feature and seemed clear on what it does, this feature can be considered a successful addition to the system.

### 6.2.6 Results: Potential impact scores

Table 6.2 shows the results from the teaching staff questionnaires when being asked about the potential impact of MarkEd overall. The last column indicates whether a change in scores was observed compared to previous students working on MarkEd. The score of 4.6 is upheld as being "very useful", which is what Andrius' first designs also received. Error-prevention dropped by 0.3 points compared to Xisen's findings; this could be attributed to the lack of indicators when saving a draft or adding a tag, and the performance issue of the page reloading on the Mark page (discussed in section 6.2.5 above). The overall net positive increase in scores indicated the additional features can be very useful to teaching staff at the School of Informatics, with comments and suggestions by users providing ample information on how to further improve.

| Area | Mean | Median | Std. Dev. | Min | Max | Change |
|---|---|---|---|---|---|---|
| Saving time | 3.6 | 4 | 1.67 | 1 | 5 | ↓ |
| Ensuring fairness | 4.2 | 4 | 0.84 | 3 | 5 | ↑ |
| Error-prevention (avoiding mis-clicks) | 3.0 | 3 | 1.41 | 1 | 5 | ↓ |
| Combining strengths of other marking tools/platforms | 3.8 | 4 | 0.84 | 3 | 5 | |
| Understanding your own progress in marking | 4.6 | 5 | 0.55 | 4 | 5 | ↑ |
| Leaving "to-do" notes for future work | 4.0 | 4 | 1.00 | 3 | 5 | |
| Ensuring high quality feedback | 3.6 | 4 | 0.55 | 3 | 4 | |
| Asking for assistance from (co-)lecturers or TAs | 4.2 | 4 | 0.84 | 3 | 5 | ↑ |

Table 6.2: Potential impact of MarkEd overall on various areas of marking

## 6.3 SUS [15] scores across all user groups

Table 6.3 displays the summary of SUS [15] scores obtained from teaching staff and students. A detailed breakdown of scores given by every single participant can be found in appendix C. Due to a low response rate by academics and markers, these two user groups were combined into one. This corresponds to an evaluation of the usability of the student pages versus the marker pages as perceived by their respective target groups.

| Role | Samples | Mean | Median | Std. Dev. | Min | Max |
|---|---|---|---|---|---|---|
| Teaching staff | 5 | 76.5 | 80 | 12.20 | 55.0 | 85 |
| Students | 21 | 81.3 | 80 | 10.94 | 52.5 | 100 |
| All | 26 | 80.4 | 80 | 11.11 | 52.5 | 100 |

Table 6.3: SUS [15] scores summary

The SUS score of 80.4 is overall a very high score in terms of perceived usability and can be thought of as being in the top 25% of products in terms of usability [11]. This score can be translated into a word descriptor of “good”. No particular skew of the data is observed with the median being very close to the mean at 80. A standard deviation of 11.11 indicates the spread of data still falls within the bounds of an “ok” system while remaining “acceptable” to users. Finally, a discrepancy between the two user groups and thus their respective interfaces can be seen. One reason for this could be the fact that the student pages overall had fewer bugs, creating an overall better user experience. Students also tended to give more positive, general comments and teaching staff more detailed, and more critical comments. This could be attributed to a professional attitude towards filling out a questionnaire and their experience with other similar tools.

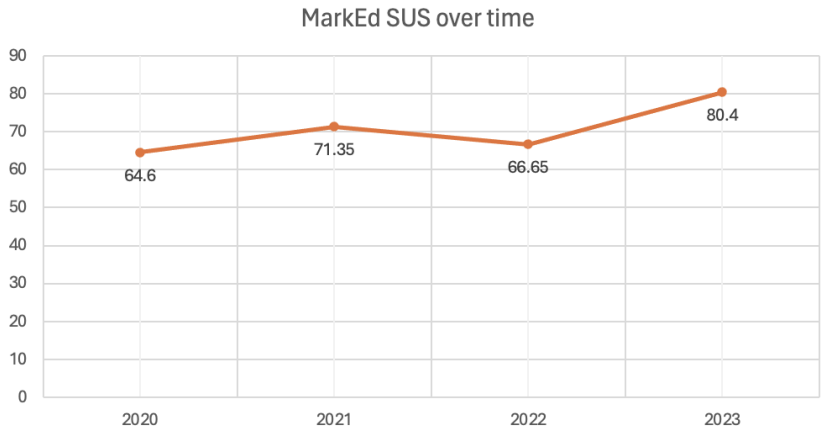



Figure 6.2: MarkEd SUS [15] scores over time

Looking at figure 6.2, there has been a 13.8 point increase in the SUS score since the last two students Xisen and Xiaofei worked on MarkEd and a 15.8 point increase since the first student Andrius. The dip from 2021 to 2022 can most likely be explained by the transition from a design prototype in Figma [35] to an actual program. Although the user has a much wider range of functionality to choose from, the higher complexity of the system now means that unexpected behaviour can also creep in as result [76]. One can see an overall positive trend in the scores. An explanation for this might be that previously functionality had only been designed and accessible in Figma [35] but after seeing the newest instalment of MarkEd on a website users felt the tool was more viable as a system they would actually use.

In summary, these good usability results, alongside all findings presented in this chapter, give a comprehensive overview of the state of MarkEd from the perspective of users in 2024.

# Chapter 7

# Concluding discussion

This chapter reviews the project as a whole, summarising achievements and providing a discussion of limitations as well as skills learned throughout. Finally, suggestions for future work and a conclusion summarising the project are presented.

## 7.1 Review

This project has answered the following research questions:

**RQ 1** - What features are missing from the current system?

Answered in chapter 4. This was done by taking the work of all 5 prior students into consideration. All the remaining 62 functional requirements were compiled into a list which can be found in appendix D.

**RQ 2** - Which features should be prioritised?

Answered in chapter 4. An adapted ICE prioritisation framework [17] was used to gain an overview of which requirements had high, medium and low priority. The tables 4.2 and 4.3 show the highest priority requirements that were subsequently implemented.

**RQ 3** - What additional steps need to be taken in order to effectively develop the new features?

Discussed in sections 5.1 and 5.4. Additional steps needed to be taken when porting to a new machine while dealing with a lack of setup instructions and database files. The system needed to be made available on the internet so that user studies could take place.

**RQ 4** - How are the new changes perceived by target user groups in terms of usability and potential impact on the marking process (for teaching staff only)?

Addressed in chapter 6. The system was rated as having the most potential impact on "understanding your own progress in marking" and "ensuring fairness". In terms of perceived usability, this project received an average SUS [15] of 80.4 across all user groups.

### 7.1.1 Achievements

The main success of this project is the positive reception to the new changes added to MarkEd. Average SUS [15] scores for user groups in the range of 76.5 to 80.4 indicate a "good", close to "excellent" system in terms of perceived usability. This can chiefly be attributed to the simplicity and clean interface for students, as seen from high scores in section 6.3. Moreover, academics as well as markers gave positive reviews on most of the new features with the system as a whole being rated as "very helpful" for "understanding your own progress", "ensuring fairness" and "asking for assistance" while marking. There is also a positive trend in user responses over the years. These results inspire confidence that MarkEd could be used in the future to make marking and moderation easier for teaching staff while maintaining the quality of feedback.

### 7.1.2 Limitations

The configuration of the Django [28] project was modified to make it more secure [92]. However, changing the Django project settings to `DEBUG = False` meant that static files were not served automatically anymore [27]. For example, when a student uploaded a file as their assignment submission to the server, Django could not find that file until the entire project was restarted, even though it had been saved. I started to create a WSGI application [29] out of the project and use the package WhiteNoise [114] to resolve this issue, but due to time constraints and my user study already being released three weeks later than planned, I was not able to finish this step. For the user study with students this issue was not a direct problem, as the tasks had all been set up without the need for viewing newly uploaded files. However, this might have been the cause of some of the problems academics and markers faced in their user studies (see chapter 6 for more detailed discussion of the results). However, for the purposes of users being able to view and use the website for themselves, the vast majority of functionality was not impacted.

### 7.1.3 Challenges

The largest challenges were project management and implementation.

This project required many strands of work to progress simultaneously. At one point implementation for iteration 2 had to be completed while integrating iteration 1 changes with my colleague, at the same time as designing a user study for these changes as well as communicating with university technicians to deploy the changes on the server. The best way to tackle this was focusing on one specific task while not losing the sight of the project as a whole, which is an approach I will definitely use in my future work.

### 7.1.4 Skills acquired

Through this project I was able to greatly develop my web development skills, particularly considering my prior inexperience in this area. This included understanding the Model-View-Controller pattern used in Django and how the client-side execution

of JavaScript code fits into this; storing state server-side; and handling asynchronous requests. This is described in detail in chapter 5 as part of the implementation process.

Finally, this project entailed no less than three separate user studies. From advertising the study via email and in-person to find participants, to moderating group discussions and interviewing users, I gained many new research and communication skills that I am certain will be useful in my future work. I feel that I have a solid understanding of NVivo [4] as a tool for thematic analysis, but would also have liked to have time to look into matrix coding queries [61] as this is a feature that I could not try out due to time constraints but would help gain a lot more insight into prevalence of themes among markers vs academics.

## 7.2 Implications for future work

Based on the features added over the duration of this project, there are still several areas for improvement.

The timer requires additional work, as some flawed logic in the implementation has a disproportionately negative impact on usability. Better implementation in this area would involve more advanced asynchronous requests [106] to reduce unnecessary refreshes and to make the time calculations more precise.

There are also five remaining medium priority requirements (see appendix D) that could be considered for the future. These are concerned with functionality for the course organisers to assign work, review marks, the necessity for user guidance or tutorials and lastly customisable tables, enabling users to sort entries by any chosen column.

The results from the user studies covered in chapter 6 demonstrate a high appreciation for the usability and potential of MarkEd. Future work can definitely shift into refining details with a heavy focus on software testing. This would be especially important as more students work on the system over time with rushed integration, increasing the risk of incurring technical debt. Testing and refactoring would hopefully help to systematically uncover bugs and prevent fixes from introducing new problems in turn [76].

## 7.3 Conclusion

While Edinburgh is already making use of a number of platforms to assist academic staff in streamlining their marking, a comprehensive system is still lacking. The aim of this project was therefore to build on work carried out by previous students to develop MarkEd further and bring it closer to a fully functional tool that incorporates strategies developed from existing research to ensure fair and efficient marking and gives teaching staff the flexibility needed to support a wide variety of assessments. As the target user groups continue to show interest and give high usability scores to the concept and current version of the system, this project has demonstrated that MarkEd could be extremely useful for the School of Informatics at the University of Edinburgh as the tool of choice for marking, moderation and feedback.

# Bibliography


[1] URL: https://www.nngroup.com/articles/success-rate-the-simplest-usability-metric/.

[2] In: (). URL: https://www.tandfonline.com/doi/epdf/10.1080/01449290310001592587?needAccess=true.

[3] "2022 Review: What's the state of the VLE market in UK higher education?" In: *Neil Mosley* (2023). URL: https://www.neilmosley.com/blog/2022-review-whats-the-state-of-the-vle-market-in-uk-higher-education.

[4] *About NVivo*. URL: https://help-nv.qsrinternational.com/20/win/Content/about-nvivo/about-nvivo.htm (visited on 04/03/2024).

[5] *Aggregation*. URL: https://docs.djangoproject.com/en/5.0/topics/db/aggregation/ (visited on 04/03/2024).

[6] Kirsti M Ala-Mutka. "A Survey of Automated Assessment Approaches for Programming Assignments". In: *Computer Science Education* (2005). DOI: 10.1080/08993400500150747.

[7] Tantoluwa Heritage Alabi. *Python requirements.txt Explained*. URL: https://www.freecodecamp.org/news/python-requirementstxt-explained/.

[8] Cristina Adriana Alexandru et al. "An Online Marking, Feedback and Moderation Tool for Computer Science Higher Education Courses". In: *Proceedings of the 7th Conference on Computing Education Practice*. 2023. DOI: https://doi.org/10.1145/3573260.3573261.

[9] Obead Alhadreti. "Assessing Academics' Perceptions of Blackboard Usability Using SUS and CSUQ: A Case Study during the COVID-19 Pandemic". In: *International Journal of Human–Computer Interaction* 37.11 (2020). DOI: https://doi.org/10.1080/10447318.2020.1861766.

[10] *Azure*. URL: https://azure.microsoft.com/en-gb (visited on 04/03/2024).

[11] Aaron Bangor, Philip T. Kortum, and James T. Miller. "An Empirical Evaluation of the System Usability Scale". In: *International Journal of Human–Computer Interaction* 24.6 (2008), pp. 574–594. DOI: 10.1080/10447310802205776.

[12] *Beacon API*. 2024. URL: https://developer.mozilla.org/en-US/docs/Web/API/Beacon_API.

[13] URL: https://getbootstrap.com/ (visited on 04/03/2024).

[14] Virginia Braun and Victoria Clarke. *Using Thematic Analysis in Psychology*. London: SAGE Publications Ltd, 2019.

[15] John Brooke. "SUS: A quick and dirty usability scale". In: *Usability Eval. Ind.* 189 (Nov. 1995).

[16] John Brooke. “SUS: A Retrospective”. In: *Journal of User Experience* 8 (2 2013), pp. 29–40. URL: https://uxpajournal.org/sus-a-retrospective/.

[17] M. Brown and S. Ellis. *Hacking Growth: How Today’s Fastest-Growing Companies Drive Breakout Success*. Ebury Publishing, 2017. ISBN: 978-0-7535-4538-6.

[18] Tom Browne, Martin Jenkins, and Richard Walker. “A longitudinal perspective regarding the use of VLEs by higher education institutions in the United Kingdom”. In: *Interactive Learning Environments* 14.2 (2006), pp. 177–192. DOI: 10.1080/10494820600852795.

[19] Bella Bruce and R. Lyons. *Universal Methods of Design: 100 Ways to Research Complex Problems, Develop Innovative Ideas, and Design Effective Solutions*. Beverly, MA: Rockport Publishers, 2012.

[20] *Button group*. URL: https://getbootstrap.com/docs/5.0/components/button-group/ (visited on 04/03/2024).

[21] W harton Cathleen John Rieman Clayton Lewis Peter Polson. “The Cognitive Walkthrough: A Practitioner’s Guide” in Usability Inspection Methods”. In: *New York: John Wiley and Sons* (1994).

[22] *Chart.js*. URL: https://www.chartjs.org/ (visited on 04/03/2024).

[23] Sun Wai Chung Chris. *Incrementing the MarkEd online marking tool with support for fairer and more efficient marking*. 2021.

[24] *CodeGrade*. URL: https://www.codegrade.com/ (visited on 04/03/2024).

[25] *ColorSpace*. colour pallete generator. URL: https://mycolor.space/?hex=%23454574&sub=1 (visited on 04/03/2024).

[26] Gennaro Cuofano. *ICE Scoring Model In A Nutshell*. FourWeekMBA. Oct. 9, 2023. URL: https://fourweekmba.com/ice-scoring-model/ (visited on 01/12/2024).

[27] *Deployment Checklist*. URL: https://docs.djangoproject.com/en/5.0/howto/deployment/checklist/#critical-settings (visited on 04/03/2024).

[28] *Django*. URL: https://www.djangoproject.com/ (visited on 04/03/2024).

[29] Django Documentation. *How to deploy with WSGI*. URL: https://docs.djangoproject.com/en/5.0/howto/deployment/wsgi/ (visited on 04/03/2024).

[30] *eduroam*. 2022. URL: https://www.ed.ac.uk/information-services/computing/desktop-personal/wifi-networking/jrs-eduroam.

[31] *Express*. URL: https://expressjs.com/ (visited on 04/03/2024).

[32] *Extended Common Marking Scheme*. 2021. URL: https://www.ed.ac.uk/timetabling-examinations/exams/regulations/common-marking-scheme.

[33] Maggie Feinberg. *The Problems with Turnitin*. 2021. URL: https://highlandpiper-sc.com/1897/opinion/the-problems-with-turnitin/.

[34] Stephen Few. “Data Visualisation Past, Present, and Future”. In: *Perceptual Edge* (2007).

[35] *Figma*. URL: https://www.figma.com/ (visited on 04/03/2024).

[36] *FortiClient VPN*. 2020. URL: https://www.ed.ac.uk/information-services/computing/desktop-personal/vpn/forticlient-vpn.

[37] *Foundation Framework*. URL: https://get.foundation/ (visited on 04/03/2024).
[38] Itamar Gilad. *Idea Prioritization With ICE and The Confidence Meter*. Apr. 18, 2018. URL: https://itamargilad.com/the-tool-that-will-help-you-choose-better-product-ideas/ (visited on 01/12/2024).
[39] Andrius Girdzius. *MarkEd: An Online Marking Tool for the University of Edinburgh School of Informatics*. 2020.
[40] *Gradescope*. URL: https://www.turnitin.com/products/gradescope/ (visited on 04/03/2024).
[41] Katherine Haan. *Best Learning Management Systems (LMS) Of 2024*. 2024.
[42] Tom Harrington. *The University of Edinburgh fares poorly in the latest National Student Survey*. 2023.
[43] AQ Haviv et al. *Web Developer's Reference Guide: A One-Stop Guide to the Essentials of Web Development Including Popular Frameworks Such As JQuery, Bootstrap, AngularJS, and Node.js*. Birmingham: Packt Publishing, Limited, 2016.
[44] "How to Conduct User Interviews". In: *nteraction Design Foundation* (2022). URL: https://www.interaction-design.org/literature/article/how-to-conduct-user-interviews#what_is_a_user_interview?-0.
[45] *How to Create Stopwatch using HTML CSS and JavaScript ?* 2023. URL: https://www.geeksforgeeks.org/how-to-create-stopwatch-using-html-css-and-javascript/ (visited on 04/03/2024).
[46] *How we prioritise bug fixes on GOV.UK - Inside GOV.UK*. Oct. 25, 2019. URL: https://insidegovuk.blog.gov.uk/2019/10/25/how-we-prioritise-bug-fixes-on-gov-uk/ (visited on 01/12/2024).
[47] Interaction Design Foundation. *How to Do a Thematic Analysis of User Interviews*. URL: https://www.interaction-design.org/literature/article/how-to-do-a-thematic-analysis-of-user-interviews.
[48] Interaction Design Foundation. *Qualitative Research*. URL: https://www.interaction-design.org/literature/topics/qualitative-research.
[49] *JavaScript*. URL: https://developer.mozilla.org/en-US/docs/Web/JavaScript (visited on 04/03/2024).
[50] Yuting Jiang. *Design of a Learning Analytics Dashboard for the School of Informatics of the University of Edinburgh*. 2023.
[51] Ankur Joshil et al. "Likert Scale: Explored and Explained". In: *British Journal of Applied Science Technology* (2015). DOI: 10.9734/BJAST/2015/14975.
[52] Tejas Kachare. *Databases for Web Developers*. URL: https://tejas-kachare18.medium.com/databases-for-web-developers-bc57802c073d (visited on 04/09/2024).
[53] J. Karlsson and K. Ryan. "A cost-value approach for prioritizing requirements". In: *IEEE Software* 14.5 (1997), pp. 67–74. DOI: 10.1109/52.605933.
[54] Amy J. Ko. *How to evaluate analytically*. URL: https://faculty.washington.edu/ajko/books/design-methods/analytical.
[55] Rachel Krause. *Consistency and Standards*. 2021. URL: https://www.nngroup.com/articles/consistency-and-standards/.
[56] *Learn Ultra*. URL: https://www.anthology.com/learn-ultra (visited on 04/03/2024).

[57] Christine Lee. *How Much Do You Know About Turnitin's Content?* 2019. URL: https://www.turnitin.com/blog/how-much-do-you-know-about-turnitins-content.

[58] Al Mansoori et al. "Faculty Perceptions of Blackboard Learn as the Main Platform for Teaching and Learning". In: *2022 International Arab Conference on Information Technology (ACIT)*. 2022, pp. 1–7. DOI: 10.1109/ACIT57182.2022.9994211.

[59] *mariaDB*. 2018-present. URL: https://mariadb.org/.

[60] Jelena Marinkovic. *Multivariate Statistics*. URL: https://link.springer.com/referenceworkentry/10.1007/978-1-4020-5614-7_2264#:~:text=Definition,between%20variables%20(Kendall%201957).

[61] *Matrix Coding Query*. URL: https://help-nv.qsrinternational.com/14/mac/Content/queries/matrix-coding-query.htm (visited on 04/03/2024).

[62] Sean McBride. *RICE: Simple prioritization for product managers*. The Intercom Blog. Jan. 5, 2018. URL: https://www.intercom.com/blog/rice-simple-prioritization-for-product-managers/ (visited on 01/12/2024).

[63] *Microsoft Forms*. URL: https://www.microsoft.com/en-gb/microsoft-365/online-surveys-polls-quizzes (visited on 04/03/2024).

[64] *Microsoft PowerPoint*. URL: https://www.microsoft.com/en-gb/microsoft-365/powerpoint (visited on 04/03/2024).

[65] *Modal*. URL: https://getbootstrap.com/docs/5.0/components/modal/ (visited on 04/03/2024).

[66] *MongoDB*. URL: https://www.mongodb.com/ (visited on 04/03/2024).

[67] *Moodle*. URL: https://moodle.org/ (visited on 04/03/2024).

[68] Marek Muszynski. "Attention checks and how to use them: Review and practical recommendations". In: *Ask Research and Methods* 32 (Dec. 2023), pp. 3–38. DOI: 10.18061/ask.v32i1.0001.

[69] URL: https://www.mysql.com/ (visited on 04/03/2024).

[70] *MySQL DB-API Drivers*. URL: https://docs.djangoproject.com/en/5.0/ref/databases/#mysql-db-api-drivers.

[71] *MySQL Workbench*. URL: https://www.mysql.com/products/workbench/ (visited on 04/03/2024).

[72] Albert Newell and Herbert A. Simon. *Human Problem Solving*. Englewood Cliffs, N.J.: Prentice Hall, 1972.

[73] Jakob Nielsen. "Chapter 6 - Usability Testing". In: *Usability Engineering* (1993). DOI: https://doi-org.eux.idm.oclc.org/10.1016/B978-0-08-052029-2.50009-7.

[74] Jakob Nielsen. *Ten Usability Heuristics*. 2024. URL: https://www.nngroup.com/articles/ten-usability-heuristics/.

[75] *OpenVPN - How and Why*. 2022. URL: https://computing.help.inf.ed.ac.uk/openvpn.

[76] Mauro Pezze and Michal Young. *Software Testing and Analysis: Process, Principles, and Techniques*. Wiley, 2007.

[77] *Popovers*. URL: https://getbootstrap.com/docs/5.0/components/popovers/ (visited on 04/03/2024).

[78] *PostgreSQL*. URL: https://www.postgresql.org/ (visited on 04/03/2024).
[79] *Programming Assignments*. URL: https://help.gradescope.com/article/ujutnle52h-instructor-assignment-programming.
[80] *Project stakeholders: Who are they and why are they important?* URL: https://www.teamwork.com/project-management-guide/project-stakeholders/.
[81] *Python 3.8*. URL: https://docs.python.org/3/whatsnew/3.8.html#notable-changes-in-3-8-18 (visited on 04/03/2024).
[82] *Python 3.8.3 Backwards Compatibility*. URL: https://peps.python.org/pep-0606/#backwards-compatibility (visited on 04/03/2024).
[83] Python Imaging Library (PIL) Fork. *Pillow*. URL: https://pypi.org/project/pillow/ (visited on 04/03/2024).
[84] *qualtrics*. URL: https://www.qualtrics.com/uk/ (visited on 04/03/2024).
[85] Ramotion. *Database in Web App Development*. URL: https://www.ramotion.com/blog/database-in-web-app-development/ (visited on 04/09/2024).
[86] Nina Abdul Razzak. *Leadership on Information Technology in Education*. Springer, 2020. DOI: https://doi.org/10.1007/978-3-030-10576-1.
[87] *Restoring Data from Dump Files*. 2018-present. URL: https://mariadb.com/kb/en/restoring-data-from-dump-files/.
[88] Colin Robson. *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*. 2002.
[89] *Ruby on Rails*. URL: https://rubyonrails.org/ (visited on 04/03/2024).
[90] Jeff Sauro. "Measuring Usability with the System Usability Scale (SUS)". In: *Measuring U* (2011). URL: https://measuringu.com/sus/.
[91] Amy Schade. "Customization vs. Personalization in the User Experience". In: *Nielsen Norman Group* (2016). URL: https://www.nngroup.com/articles/customization-personalization/.
[92] Matt Segal. *Simple Django Deployment*. 2020. URL: https://mattsegal.dev/simple-django-deployment-3.html.
[93] Navin Singh. *Top 10 Backend Web Development Frameworks 2023*. URL: https://www.linkedin.com/pulse/top-10-backend-web-development-frameworks-2023-navin-singh (visited on 04/09/2024).
[94] *Spring Boot*. URL: https://spring.io/projects/spring-boot (visited on 04/03/2024).
[95] *SQLite*. URL: https://www.sqlite.org/ (visited on 04/03/2024).
[96] *stack overflow*. URL: https://stackoverflow.com/ (visited on 04/03/2024).
[97] *Student Submit Work*. URL: https://help.gradescope.com/article/ccbpppziu9-student-submit-work#submitting_code (visited on 04/03/2024).
[98] Xiaofei Sun. *MarkEd: An Online Marking Tool for the University of Edinburgh School of Informatics*. 2021.
[99] *SurveyLegend*. URL: https://www.surveylegend.com/ (visited on 04/03/2024).
[100] *tailwindcss*. URL: https://tailwindcss.com/ (visited on 04/03/2024).
[101] Sofiia Tarhonska. *Use Django to send emails with SMTP*. 2022. URL: https://opensource.com/article/22/12/django-send-emails-smtp.
[102] *The best UK universities 2024 – rankings*. 2024. URL: https://www.theguardian.com/education/ng-interactive/2023/sep/09/the-guardian-university-guide-2024-the-rankings.

[103] *The FORM element*. URL: https://www.w3.org/MarkUp/html3/forms.html (visited on 04/03/2024).

[104] *Turnitin*. URL: https://www.turnitin.co.uk/ (visited on 04/03/2024).

[105] *Turnitin Frequently Asked Questions (FAQs) for Students*. URL: https://services.anu.edu.au/information-technology/software-systems/turnitin/turnitin-frequently-asked-questions-faqs-for-0.

[106] Brennan Tymrak. *Fetching Data with AJAX and Django*. 2019-2024. URL: https://www.brennantymrak.com/articles/fetching-data-with-ajax-and-django.

[107] *Use of Color: Understanding SC 1.4.1*. 2023. URL: https://www.w3.org/TR/UNDERSTANDING-WCAG20/visual-audio-contrast-without-color.html.

[108] Kami Vaniea. *Gestalt Principles*. lecture. 2023. URL: https://opencourse.inf.ed.ac.uk/sites/default/files/https/opencourse.inf.ed.ac.uk/hci/2023/hci-2020-gestalt-principles.pdf.

[109] Kami Vaniea. *HCI Evaluation*. lecture. 2020. URL: https://opencourse.inf.ed.ac.uk/sites/default/files/https/opencourse.inf.ed.ac.uk/hci/2023/hci-2020evaluation-what-it-and-why-do-it.pdf.

[110] Kami Venea. *Interviews*. lecture. URL: https://opencourse.inf.ed.ac.uk/sites/default/files/https/opencourse.inf.ed.ac.uk/hci/2023/hci-2020interviews.pdf.

[111] Wan Mohd Hamdani bin Wan Mohd Azmi. *Incrementing the MarkEd online marking tool with support for good feedback*. 2021.

[112] Xisen Wang. *The MarkEd tool for Marking, Feedback and Moderation II*. 2021.

[113] Natalie Weiner. *All mixed up: You, me, and UI*. 2023. URL: https://www.theverge.com/23653818/spotify-shuffle-button-music-history.

[114] *WhiteNoise Documentation*. URL: https://whitenoise.readthedocs.io/en/latest/ (visited on 04/03/2024).

[115] *Window: load event*. 2023. URL: https://developer.mozilla.org/en-US/docs/Web/API/Window/load_event.

[116] *Youri Voet*. 2021. URL: https://www.timeshighered-events.com/digital-universities-week-uk-2021/agenda/speakers/1165762.

# Appendix A

# Participants' consent form

1. Consent form for text-based data collection
2. Consent form for video recording data collection

Participant number:_____

# Participant Consent Form

| | |
|---|---|
| Project title: | The MarkEd tool for Marking, Feedback and Moderation III |
| Principal investigator: | Dr Cristina Adriana Alexandru |
| Researcher collecting data: | Lea Andrusz (main researcher) |
| PI contact details: | Cristina.Alexandru@ed.ac.uk |

By participating in the study you agree that:

- I have read and understood the Participant Information Sheet for the above study, that I have had the opportunity to ask questions, and that any questions I had were answered to my satisfaction.
- My participation is voluntary, and that I can withdraw at any time without giving a reason. Withdrawing will not affect any of my rights.
- I consent to my anonymised data being used in academic publications and presentations.
- I understand that my anonymised data will be stored for the duration outlined in the Participant Information Sheet.

**Please tick yes or no for each of these statements.**

| | | | |
|---|---|---|---|
| **1.** | I allow my data to be used in future ethically approved research. | **Yes** ☐ | **No** ☐ |
| **2.** | I agree to take part in this study. | **Yes** ☐ | **No** ☐ |

| Name of person giving consent | Date dd/mm/yy | Signature |
|---|---|---|
| | | |

| Name of person taking consent | Date dd/mm/yy | Signature |
|---|---|---|
| | | |

THE UNIVERSITY of EDINBURGH informatics

Participant number:_____

# Participant Consent Form

| | |
|---|---|
| Project title: | The MarkEd tool for Marking, Feedback and |
| Principal investigator: | Dr Cristina Adriana Alexandru |
| Researcher collecting | Lea Andrusz (main researcher), Anish Thapa |
| PI contact details: | Cristina.Alexandru@ed.ac.uk |

By participating in the study you agree that:

- I have read and understood the Participant Information Sheet for the above study, that I have had the opportunity to ask questions, and that any questions I had were answered to my satisfaction.
- My participation is voluntary, and that I can withdraw at any time without giving a reason. Withdrawing will not affect any of my rights.
- I consent to my anonymised data being used in academic publications and presentations.
- I understand that my anonymised data will be stored for the duration outlined in the Participant Information Sheet.

**Please tick yes or no for each of these statements.**

| | | | |
|---|---|---|---|
| **1.** | I agree to being audio recorded. | | **Yes** / **No** |
| **2.** | I agree to being video recorded. | | **Yes** / **No** |
| **3.** | I allow my data to be used in future ethically approved research. | | **Yes** / **No** |
| **4.** | I agree to take part in this study. | | **Yes** / **No** |

| Name of person giving consent | Date<br>dd/mm/yy | Signature |
|---|---|---|
| | | |

| Name of person taking consent | Date<br>dd/mm/yy | Signature |
|---|---|---|
| | | |

THE UNIVERSITY of EDINBURGH informatics

# Appendix B

# Participants' information sheet

1. Participant information sheet for questionnaire
2. Participant information sheet for group study
3. Participant information sheet for individual study

## Participant Information Sheet

| | |
|---|---|
| Project title: | The MarkEd tool for Marking, Feedback and Moderation III |
| Principal investigator: | Dr Cristina Adriana Alexandru |
| Researcher collecting data: | Lea Andrusz (main researcher) |
| Funder (if applicable): | No |

This study was certified according to the Informatics Research Ethics Process, reference number 681259. Please take time to read the following information carefully. You should keep this page for your records.

**Who are the researchers?**

Lea Andrusz, Dr Alexandru (supervisor)

**What is the purpose of the study?**

This study aims to evaluate the usability and potential impact of a new iteration of the implemented MarkEd system. This iteration is currently being developed as part of the BSc honours project to create a tool for submissions, marking, feedback and moderation of student assignments. It aims to have a simple, intuitive interface with included elements to foster efficient and fair marking, and high-quality feedback, based on research conducted by previous students. The results will allow for us to improve the tool such that it meets users' expectations and we hope that the system will ultimately be adopted by the School of Informatics at the University of Edinburgh.

**Why have I been asked to take part?**

You are a part of one or more key user groups of this tool:

- a student
- a member of the academic staff
- a marker

within the School of Informatics. Your experience with either submitting assignments and receiving feedback or your experience with marking, giving feedback and moderation is vital for understanding how helpful MarkEd can be for you. Your

contribution is important to determine the usability for the proposed tool, which could improve the student and staff experience with assessment in the future.

**Do I have to take part?**

No – participation in this study is entirely up to you. You can withdraw from the study at any time, without giving a reason. Your rights will not be affected. If you wish to withdraw, contact the PI. We will stop using your data in any publications or presentations submitted after you have withdrawn consent. However, we will keep copies of your original consent, and of your withdrawal request.

**What will happen if I decide to take part?**

- Lea Andrusz will email you the link to the prototype, as well as a link to an online questionnaire on Microsoft Forms.
- Questions will start with those about your year of study (for students)/ level of experience with marking, feedback and moderation (academic staff). They will then revolve around tasks asked to complete within our prototype and your perception of the difficulty of these. You may be asked to upload screenshots of our prototype interface via the form to support suggestions you make.
- You may be asked to choose your preferred design out of several options. General questions about the overall usability of the prototype will be asked.
- This should take around 5-15 minutes.

**Are there any risks associated with taking part?**

There are no significant risks associated with participation.

**Are there any benefits associated with taking part?**

Direct benefits: None.

Indirect benefits: You will play an important role in developing a better system for more efficient marking and higher quality feedback in the School of Informatics.

**What will happen to the results of this study?**

The results will be shared in Lea Andrusz's BSc honours project. The results of this study may be summarised in published articles, reports and presentations. Quotes or

key findings will be anonymized: We will remove any information that could, in our assessment, allow anyone to identify you. With your consent, information can also be used for future research. Your data may be archived for a minimum of 2 years.

**Data protection and confidentiality.**

Your data will be processed in accordance with Data Protection Law. All information collected about you will be kept strictly confidential. Your data will be referred to by a unique participant number rather than by name. Your data will only be viewed by the researcher Lea Andrusz and supervisor Dr Cristina Alexandru

All electronic data will be stored on a password-protected encrypted computer or on the University's secure encrypted cloud storage services (DataShare, ownCloud, or Sharepoint) and all paper records will be stored in a locked filing cabinet in the PI's office. Your consent information will be kept separately from your responses in order to minimise risk.

**What are my data protection rights?**

The University of Edinburgh is a Data Controller for the information you provide. You have the right to access information held about you. Your right of access can be exercised in accordance Data Protection Law. You also have other rights including rights of correction, erasure and objection. For more details, including the right to lodge a complaint with the Information Commissioner's Office, please visit www.ico.org.uk. Questions, comments and requests about your personal data can also be sent to the University Data Protection Officer at dpo@ed.ac.uk.

For general information about how we use your data, go to: edin.ac/privacy-research

**Who can I contact?**

If you have any further questions about the study, please contact the lead researcher, Lea Andrusz (L.A.Andrusz@sms.ed.ac.uk)

If you wish to make a complaint about the study, please contact inf-ethics@inf.ed.ac.uk. When you contact us, please provide the study title and detail the nature of your complaint.

**Updated information.**

If the research project changes in any way, an updated Participant Information Sheet will be emailed to you by Lea Andrusz (L.A.Andrusz@sms.ed.ac.uk).

**Consent**

By proceeding with the study, I agree to all of the following statements:

- I have read and understood the above information.
- I understand that my participation is voluntary, and I can withdraw at any time.
- I consent to my anonymised data being used in academic publications and presentations.
- I allow my data to be used in future ethically approved research.

# Participant Information Sheet

| | |
|---|---|
| Project title: | The MarkEd tool for Marking, Feedback and Moderation III |
| Principal investigator: | Dr Cristina Adriana Alexandru |
| Researcher collecting data: | Lea Andrusz (main researcher), Anish Thapa |
| Funder (if applicable): | No |

This study was certified according to the Informatics Research Ethics Process, reference number 681259. Please take time to read the following information carefully. You should keep this page for your records.

**Who are the researchers?**

Lea Andrusz, Anish Thapa, Dr Alexandru (supervisor)

**What is the purpose of the study?**

This study aims to evaluate the usability and potential impact of a new iteration of the implemented MarkEd system. This iteration is currently being developed as part of the BSc honours project to create a tool for submissions, marking, feedback and moderation of student assignments. It aims to have a simple, intuitive interface with included elements to foster efficient and fair marking, and high-quality feedback, based on research conducted by previous students. The results will allow for us to improve the tool such that it meets users' expectations and we hope that the system will ultimately be adopted by the School of Informatics at the University of Edinburgh.

**Why have I been asked to take part?**

You are a part of one or more key user groups of this tool:

- a student
- a member of the academic staff
- a marker

within the School of Informatics. Your experience with either submitting assignments and receiving feedback or your experience with marking, giving feedback and moderation is vital for understanding how helpful MarkEd can be for you. Your

contribution is important to determine the usability for the proposed tool, which could improve the student and staff experience with assessment in the future.

**Do I have to take part?**

No – participation in this study is entirely up to you. You can withdraw from the study at any time, without giving a reason. Your rights will not be affected. If you wish to withdraw, contact the PI. We will stop using your data in any publications or presentations submitted after you have withdrawn consent. However, we will keep copies of your original consent, and of your withdrawal request.

**What will happen if I decide to take part?**

- You will be invited to participate to a group meeting with other people with the same student / academic staff and/or marker member role as you, organised online on Microsoft Teams at a time that is suitable to you.
- Part 1: Stakeholder walkthrough. You will be asked to complete a series of tasks using the MarkEd prototype. These tasks will be typical to your role. You will be asked to share your thoughts after you complete these tasks in a group setting.
- Part 2: Focus group. You will then participate in a group discussion on positive and negative experiences with the prototype as well as suggestions on the prototype and the potential impact of this prototype on your future work.
- Part 3: Questionnaire. Finally, we will ask you to fill in a short questionnaire about your year of study (for students)/ level of experience with marking, feedback and moderation (academic staff), and your opinions on the system's general usability (SUS questionnaire), using a Microsoft Forms form.
- The whole process should take around 30 minutes.
- Lea Andrusz will lead the discussion, while Anish Thapa and Cristina Alexandru will be taking notes. With your approval, we would also like to record the meeting.

**Are there any risks associated with taking part?**

There are no significant risks associated with participation.

**Are there any benefits associated with taking part?**

Direct benefits: none
Indirect benefits: You will play an important role in developing a better system for more efficient marking and higher quality feedback in the School of Informatics.

**What will happen to the results of this study?**

The results will be shared in Lea Andrusz's BSc honours project. The results of this study may be summarised in published articles, reports and presentations. Quotes or key findings will be anonymized: We will remove any information that could, in our assessment, allow anyone to identify you. With your consent, information can also be used for future research. Your data may be archived for a minimum of 2 years.

**Data protection and confidentiality.**

Your data will be processed in accordance with Data Protection Law. All information collected about you will be kept strictly confidential. Your data will be referred to by a unique participant number rather than by name. Your data will only be viewed by the researcher Lea Andrusz and supervisor Dr Cristina Alexandru

All electronic data will be stored on a password-protected encrypted computer or on the University's secure encrypted cloud storage services (DataShare, ownCloud, or Sharepoint) and all paper records will be stored in a locked filing cabinet in the PI's office. Your consent information will be kept separately from your responses in order to minimise risk.

**What are my data protection rights?**

The University of Edinburgh is a Data Controller for the information you provide. You have the right to access information held about you. Your right of access can be exercised in accordance Data Protection Law. You also have other rights including rights of correction, erasure and objection. For more details, including the right to lodge a complaint with the Information Commissioner's Office, please visit www.ico.org.uk. Questions, comments and requests about your personal data can also be sent to the University Data Protection Officer at dpo@ed.ac.uk.
For general information about how we use your data, go to: edin.ac/privacy-research

**Who can I contact?**

If you have any further questions about the study, please contact the lead researcher, Lea Andrusz (L.A.Andrusz@sms.ed.ac.uk)

If you wish to make a complaint about the study, please contact inf-ethics@inf.ed.ac.uk. When you contact us, please provide the study title and detail the nature of your complaint.

**Updated information.**

If the research project changes in any way, an updated Participant Information Sheet will be emailed to you by Lea Andrusz (L.A.Andrusz@sms.ed.ac.uk).

**Consent**

By proceeding with the study, I agree to all of the following statements:

- I have read and understood the above information.
- I understand that my participation is voluntary, and I can withdraw at any time.
- I consent to my anonymised data being used in academic publications and presentations.
- I allow my data to be used in future ethically approved research.

# Participant Information Sheet

| | |
|---|---|
| Project title: | The MarkEd tool for Marking, Feedback and Moderation III |
| Principal investigator: | Dr Cristina Adriana Alexandru |
| Researcher collecting data: | Lea Andrusz (main researcher), Anish Thapa |
| Funder (if applicable): | No |

This study was certified according to the Informatics Research Ethics Process, reference number 681259. Please take time to read the following information carefully. You should keep this page for your records.

**Who are the researchers?**

Lea Andrusz, Anish Thapa, Dr Alexandru (supervisor)

**What is the purpose of the study?**

This study aims to evaluate the usability and potential impact of a new iteration of the implemented MarkEd system. This iteration is currently being developed as part of the BSc honours project to create a tool for submissions, marking, feedback and moderation of student assignments. It aims to have a simple, intuitive interface with included elements to foster efficient and fair marking, and high-quality feedback, based on research conducted by previous students. The results will allow for us to improve the tool such that it meets users' expectations and we hope that the system will ultimately be adopted by the School of Informatics at the University of Edinburgh.

**Why have I been asked to take part?**

You are a part of one or more key user groups of this tool:

- a student
- a member of the academic staff
- a marker

within the School of Informatics. Your experience with either submitting assignments and receiving feedback or your experience with marking, giving feedback and moderation is vital for understanding how helpful MarkEd can be for you. Your

contribution is important to determine the usability for the proposed tool, which could improve the student and staff experience with assessment in the future.

**Do I have to take part?**

No – participation in this study is entirely up to you. You can withdraw from the study at any time, without giving a reason. Your rights will not be affected. If you wish to withdraw, contact the PI. We will stop using your data in any publications or presentations submitted after you have withdrawn consent. However, we will keep copies of your original consent, and of your withdrawal request.

**What will happen if I decide to take part?**

- You will be invited to participate to a one-to-one meeting organised online on Microsoft Teams at a time that is suitable to you.
- Part 1: Think aloud. You will be asked to complete a series of tasks using the MarkEd prototype. These tasks will be typical to your role (student / academic staff or marker). You will be asked to share your thoughts after you complete these tasks in a group setting.
- Part 2: Semi-structured interview. You will then be asked follow up questions on positive and negative experiences with the prototype and questions about the potential impact of this prototype on your future work.
- Part 3: Questionnaire. Finally, we will ask you to fill in a short questionnaire about your year of study (for students)/ level of experience with marking, feedback and moderation (academic staff), and your opinions on the system's general usability (SUS questionnaire), using a Microsoft Forms form.
- The whole process should take around 20 minutes.
- Lea Andrusz will lead the discussion, while Anish Thapa and Cristina Alexandru will be taking notes. With your approval, we would also like to record the meeting.

**Are there any risks associated with taking part?**

There are no significant risks associated with participation.

**Are there any benefits associated with taking part?**

Direct benefits: none

Indirect benefits: You will play an important role in developing a better system for more efficient marking and higher quality feedback in the School of Informatics.

**What will happen to the results of this study?**

The results will be shared in Lea Andrusz's BSc honours project. The results of this study may be summarised in published articles, reports and presentations. Quotes or key findings will be anonymized: We will remove any information that could, in our assessment, allow anyone to identify you. With your consent, information can also be used for future research. Your data may be archived for a minimum of 2 years.

**Data protection and confidentiality.**

Your data will be processed in accordance with Data Protection Law. All information collected about you will be kept strictly confidential. Your data will be referred to by a unique participant number rather than by name. Your data will only be viewed by the researcher Lea Andrusz and supervisor Dr Cristina Alexandru

All electronic data will be stored on a password-protected encrypted computer or on the University's secure encrypted cloud storage services (DataShare, ownCloud, or Sharepoint) and all paper records will be stored in a locked filing cabinet in the PI's office. Your consent information will be kept separately from your responses in order to minimise risk.

**What are my data protection rights?**

The University of Edinburgh is a Data Controller for the information you provide. You have the right to access information held about you. Your right of access can be exercised in accordance Data Protection Law. You also have other rights including rights of correction, erasure and objection. For more details, including the right to lodge a complaint with the Information Commissioner's Office, please visit www.ico.org.uk. Questions, comments and requests about your personal data can also be sent to the University Data Protection Officer at dpo@ed.ac.uk.
For general information about how we use your data, go to: edin.ac/privacy-research

**Who can I contact?**

If you have any further questions about the study, please contact the lead researcher, Lea Andrusz (L.A.Andrusz@sms.ed.ac.uk)
If you wish to make a complaint about the study, please contact

inf-ethics@inf.ed.ac.uk. When you contact us, please provide the study title and detail the nature of your complaint.

**Updated information.**

If the research project changes in any way, an updated Participant Information Sheet will be emailed to you by Lea Andrusz (L.A.Andrusz@sms.ed.ac.uk).

**Consent**

By proceeding with the study, I agree to all of the following statements:

- I have read and understood the above information.
- I understand that my participation is voluntary, and I can withdraw at any time.
- I consent to my anonymised data being used in academic publications and presentations.
- I allow my data to be used in future ethically approved research.

# Appendix C

# User study details

## Participant details

| Student | Year |
|---|---|
| S1 | Year 4 |
| S2 | Year 4 |
| S3 | Year 4 |
| S4 | Year 4 |
| S5 | Year 4 |
| S6 | Year 4 |
| S7 | Year 4 |
| S8 | Year 4 |
| S9 | Year 4 |
| S10 | Year 4 |
| S11 | Year 4 |
| S12 | PG Research |
| S13 | MSc |
| S14 | PG Research |
| S15 | Year 4 |
| S16 | Year 4 |
| S17 | Year 4 |
| S18 | Year 4 |
| S19 | Year 1 |
| S20 | Year 5 |
| S21 | PG Research |
| S22 | Year 1 |
| S23 | Year 4 |
| S24 | Year 5 |

Table C.1: Student participant details

| My ID | Involvement in Course Organization | Involvement in Marking Duties |
|---|---|---|
| A1 | Over 5 years | 4 years or more |
| A2 | 3-5 years | 4 years or more |
| A3 | unknown | unknown |
| A4 | unknown | unknown |
| A5 | unknown | unknown |
| M1 | n/a | 6 months or less |
| M2 | n/a | 2-3 years |
| M3 | n/a | 6 months or less |
| M4 | unknown | unknown |
| M5 | unknown | unknown |

Table C.2: Academic and Marker participant details

# Student Questionnaire

## Questions

# MarkEd: the tool for marking, moderation and feedback - STUDENT questionnaire

Questionnaire to help evaluate work done on the student interfaces for the tool.

* Required

## Introduction

1

***Please take a few minutes to read the following participant information sheet:***
https://uoe-my.sharepoint.com/:b:/g/personal/s1953043_ed_ac_uk/ESLyrKhLE2lDiqmaDPg59GsBdMFluCszQcoxj50mbdDddw

Please select "yes" if you agree to **ALL** of the statements below:

- I have read and understood the Participant Information Sheet for the above study, that I have had the opportunity to ask questions, and that any questions I had were answered to my satisfaction.
- My participation is voluntary, and that I can withdraw at any time without giving a reason. Withdrawing will not affect any of my rights.
- I consent to my anonymised data being used in academic publications and presentations.
- I understand that my anonymised data will be stored for the duration outlined in the Participant Information Sheet.
- I allow my data to be used in future ethically approved research.
- I agree to take part in this study.

*

◯ Yes

◯ No

2

Are you an undergraduate or postgraduate taught student? *

◯ Yes

◯ No

3

What year are you in? *

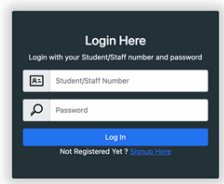
- ◯ year 1
- ◯ year 2
- ◯ year 3
- ◯ year 4
- ◯ year 5
- ◯ MSc

4

**Follow the link: http://104.40.244.158/login/**

Are you able to view the **login** page for MarkEd?

The image shows what it should look like. *



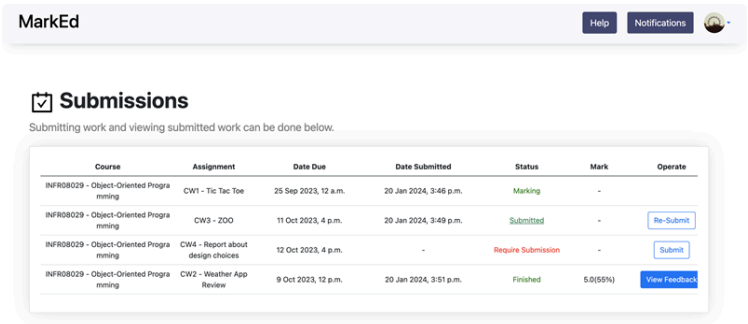

- ◯ Yes
- ◯ No

5

Please log in to MarkEd with the following credentials:

Student number = **2020202**
and Password = **study333**

Have you managed to do this successfully and do you see a **home page** similar to the one shown in the image? *




- ◯ Yes
- ◯ No

6

If you were unable to complete this page, please explain **why**. *

7

Please enter the due date for assignment "CW4 - Report about design choices" for the "Object-Oriented Programming" course. *

8

Which of these statements do you most agree with? *

- ◯ The order of the assignments listed on the home page is **intuitive**.
- ◯ I **don't care** about the order of the assignments listed on the home page.
- ◯ The order of the assignments listed on the home page is **unintuitive / bothersome**.

9

If you found the order **unintuitive** in the question above, please indicate the default ordering of assignments you would prefer. *

- ◯ In order of due date (earliest first)
- ◯ In order of due date (latest first)
- ◯ In order of status ('un-submitted' assignments first, submitted, marking, then 'finished' last)
- ◯ Other

10

Please rate the **content**, **presentation** and **interaction** for the **home page**. *

**content** could be the information displayed. Is it as expected? Is anything missing?
**presentation** could be page layout, visual hierarchy, colour/font styles
**interaction** could navigation and any interactive elements (buttons, tooltips)

| | Poor | Fair | Good | Very good |
|---|---|---|---|---|
| content | ◯ | ◯ | ◯ | ◯ |
| presentation | ◯ | ◯ | ◯ | ◯ |
| interaction | ◯ | ◯ | ◯ | ◯ |

11

Do you have any suggestions for this page?

Especially for any of the content, presentation, layout that you rated as "poor" or "fair".

## Task 1/4 - Submit file

Please submit any pdf (can be blank) for the assignment "CW4 - Report about design choices" for the "Object-Oriented Programming" course.

12

Have you managed to do this successfully (i.e. have you received a confirmation message at the top of the page)? *

- ◯ Yes
- ◯ No

13

If you were unable to complete this task, please explain **why**. *

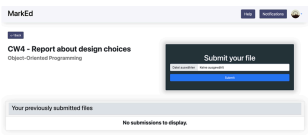
14

What was the name of the file you just submitted? *

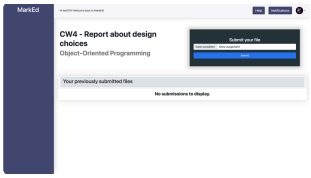
15

Which of these three page designs for submitting files do you prefer? *

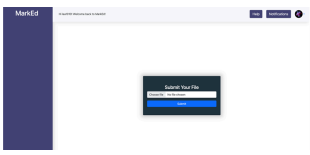
Zoom in to see details.




◯ Option 1 (view submitted file with wider screen space)




◯ Option 2 (view submitted file with narrower screen space)




◯ Option 3 (no option to view the submitted file)

16

Please give a reason for your choice. *

## Task 2/4 - View submission

Please view the previously submitted file for the assignment "CW1 - Tic Tac Toe" for the "Object-Oriented Programming" course.

17

Have you managed to do this successfully? *

- ◯ Yes
- ◯ No

18

If you were unable to complete this task, please explain **why**. *

19

How did you achieve this? *

- ◯ "Re-submit" button
- ◯ "Submitted" status link
- ◯ Other

20

Do you have any further suggestions regarding this task? *

21

Please rate the **content**, **presentation** and **interaction** for the **submit / view submission page**. *

**content** could be the information displayed. Is it as expected? Is anything missing?
**presentation** could be page layout, visual hierarchy, colour/font styles
**interaction** could navigation and any interactive elements (buttons, tooltips)

| | Poor | Fair | Good | Very good |
|---|---|---|---|---|
| content | ◯ | ◯ | ◯ | ◯ |
| presentation | ◯ | ◯ | ◯ | ◯ |
| interaction | ◯ | ◯ | ◯ | ◯ |

22

Do you have any suggestions for this page?

Especially for any of the content, presentation, layout that you rated as "poor" or "fair".

## Task 3/4 - View feedback

Please view your feedback for assignment "CW2 - Weather App Review" for the "Object-Oriented Programming" course.

23

Have you managed to do this successfully? *

- ◯ Yes
- ◯ No

24

If you were unable to complete this task, please explain **why**. *

25

How many attempts were submitted for this assignment? *

26

Please rate the **content**, **presentation** and **interaction** for the **feedback page**. *

**content** could be the information displayed. Is it as expected? Is anything missing?
**presentation** could be page layout, visual hierarchy, colour/font styles
**interaction** could navigation and any interactive elements (buttons, tooltips)

| | Poor | Fair | Good | Very good |
|---|---|---|---|---|
| content | ◯ | ◯ | ◯ | ◯ |
| presentation | ◯ | ◯ | ◯ | ◯ |
| interaction | ◯ | ◯ | ◯ | ◯ |

27

Do you have any suggestions for this page?

Especially for any of the content, presentation, layout that you rated as "poor" or "fair".

## Task 4/4 - Notifications

Please view all your notifications.

28

Have you managed to do this successfully? *

- ◯ Yes
- ◯ No

29

If you were unable to complete this task, please explain **why**. *

30

How did you achieve this? *

Please provide a short series of steps you took to get to your notifications.

31

Do you have any further suggestions regarding this task?

32

Please rate the **content**, **presentation** and **interaction** for the **notification page**. *

**content** could be the information displayed. Is it as expected? Is anything missing?
**presentation** could be page layout, visual hierarchy, colour/font styles
**interaction** could navigation and any interactive elements (buttons, tooltips)

| | Poor | Fair | Good | Very good |
|---|---|---|---|---|
| content | ◯ | ◯ | ◯ | ◯ |
| presentation | ◯ | ◯ | ◯ | ◯ |
| interaction | ◯ | ◯ | ◯ | ◯ |

Do you have any suggestions for this page?

Especially for any of the content, presentation, layout that you rated as "poor" or "fair".

## General questions

34

How have you been navigating back to the home page between tasks? *


Browser 'back' button


'MarkEd' in the top left corner


Interface 'back' button

35

Please indicate how you would like to receive notifications for the following events. *

| | email | in-app | BOTH (email and in-app) | NO notification |
|---|---|---|---|---|
| First time log in welcome. | ◯ | ◯ | ◯ | ◯ |
| You have been added to a new course. | ◯ | ◯ | ◯ | ◯ |
| A new assignment has been added. | ◯ | ◯ | ◯ | ◯ |
| You have made a new submission. | ◯ | ◯ | ◯ | ◯ |
| Marks have been released. | ◯ | ◯ | ◯ | ◯ |

36

Is there anything you'd like to add regarding notifications for MarkEd?

**For example:**
Would you like to receive notifications for anything else?

## Last section - System Usability Scale (SUS)

A standard questionnaire used to assess the usability of a product.

37

If you were given the choice as to what system to use for your assignments, please indicate your opinion regarding these 10 statements. *

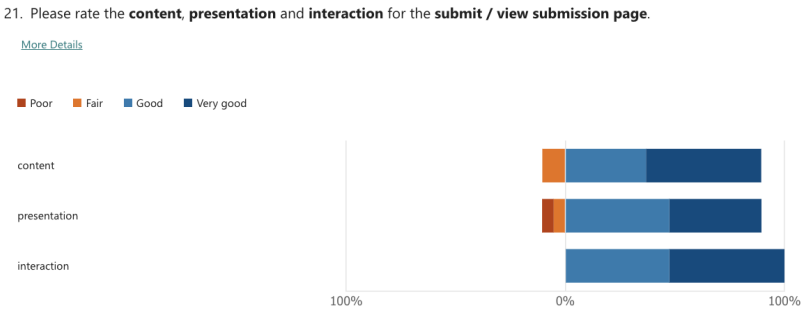
| | STRONGLY disagree | disagree | neutral | agree | STRONGLY agree |
|---|---|---|---|---|---|
| I think that I would like to use this system (MarkEd) frequently. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I found the system unnecessarily complex. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I thought the system was easy to use. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I think that I would need the support of a technical person to be able to use this system. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I found the various functions in this system were well integrated. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I thought there was too much inconsistency in this system. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I would imagine that most students would learn to use this system very quickly. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I found the system very cumbersome to use. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I felt very confident using the system. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I needed to learn a lot of things before I could get going with this system. | ◯ | ◯ | ◯ | ◯ | ◯ |

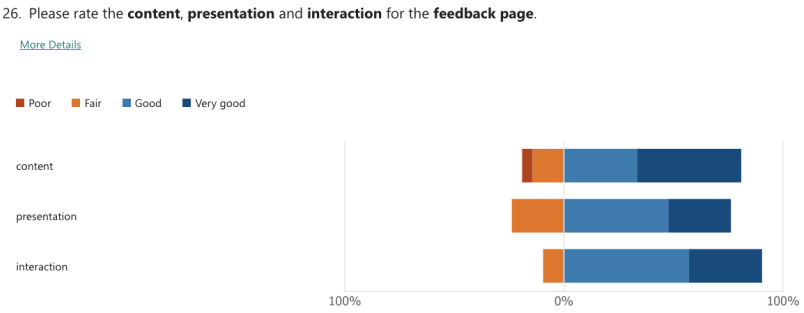
## Thank you

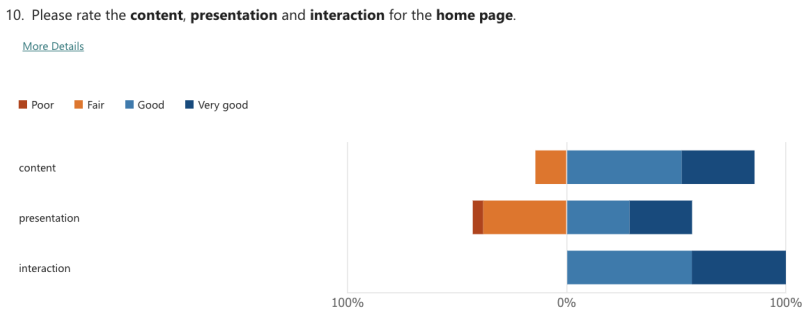
38

If you are willing to participate in a follow-up interview, please leave your email address below.

## Content, presentation and interaction results




## Stakeholder walkthrough presentation including tasks and questions

Please see the next page.

# MarkEd

The tool for marking, feedback and moderation

Stakeholder walkthrough 11 March 2024

1

Participant Information Sheet + Consent Forms
LINK:
https://forms.office.com/Pages/ResponsePage.aspx?id=sAafLmkWiUWHiRCeaTTcYc_WwQniee91KkTqFM_90YpxUMDg3NQplMDZBRkM2TTNHTFgxS1NLTkMwWS4u

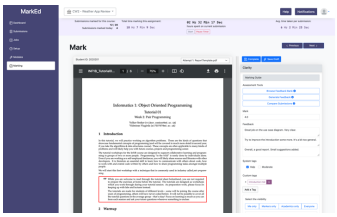
# MarkEd

The tool for marking, feedback and moderation

Stakeholder walkthrough 11 March 2024

2

## MarkEd – quick introduction



- 1 platform
- More **flexibility**
- Incorporating strategies for **fair** and **efficient** marking

Blackboard LEARN ULTRA

gradescope

turnitin

codegrade

3

## Today's plan

1. Complete 4 tasks using MarkEd. Share your thoughts with the group after each task. [25 min]
2. General comments / suggestions [5 min]
3. Standardised questionnaire on usability and impact of the tool. [10 min]

Mark a submission → View the dashboard → Moderate

4

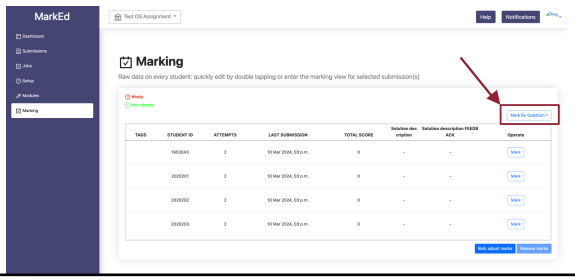
## Task 0 part 1 – log in

- MarkEd LINK: http://104.40.244.158/login/
- Log in using the credentials in the table
- Select the correct assignment for the course "Operating Systems"
- Go to the "marking" tab in the sidebar

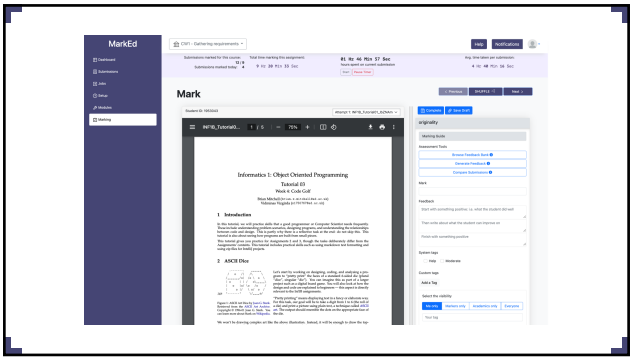
| Staff number | password | Assignment name |
|---|---|---|
| 603 | aca603 | CW1 Copy 1 |
| 604 | aca604 | CW1 Copy 2 |
| 605 | aca605 | CW1 Copy 3 |
| 606 | aca606 | CW1 Copy 4 |
| 607 | aca607 | CW1 Copy 5 |
| 608 | aca608 | CW1 Copy 6 |

5

## Task 0 part 2 - select 'mark by question'




6

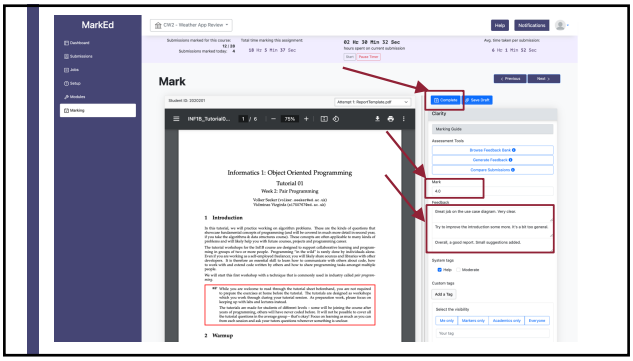



7

## Task 0 part 3 – mark a submission

1. Start marking by filling out the "**Mark**" and "**Feedback**" fields in the panel on the right side
   1. *Give a mark between 0 and 5*
2. Save your marking using the "Complete" button

8

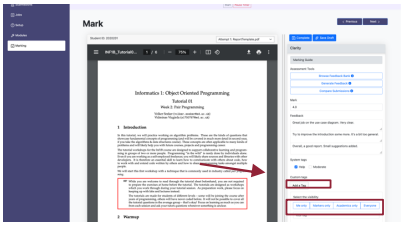



9

## Task 1 – custom tags

1. Add two custom tags for question "solution description"
   1. *One tag only visible to you (for example: re-check intro)*
   2. *Add a tag visible to everyone (short note about the submission for all staff to see)*
2. Delete the first tag

10



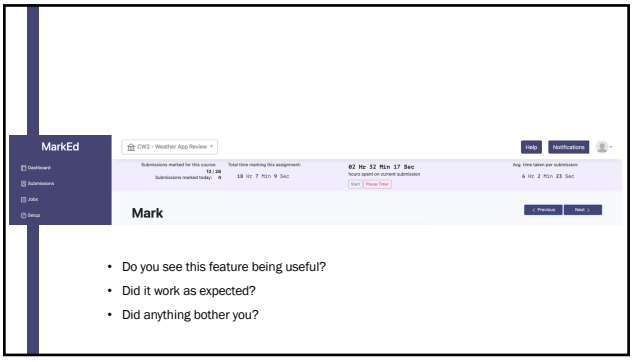

- Do you see this feature being useful?
- Did it work as expected?
- Did anything bother you?

11

## Task 2 – timer

- Have a look at the timer and statistics (at the top of the page)
- Pause and start the timer
- Read through the statistics generated

12



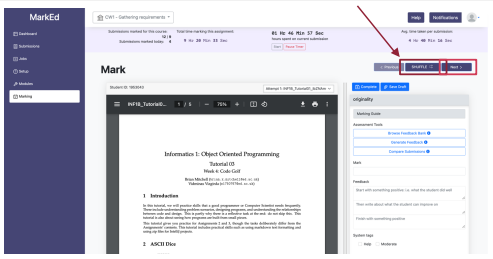

- Do you see this feature being useful?
- Did it work as expected?
- Did anything bother you?

13

## Task 3 – Shuffle by question

1. Turn on shuffle
2. Navigate to the next submission (do this three times).
3. Go back one submission
4. Turn off shuffle

14



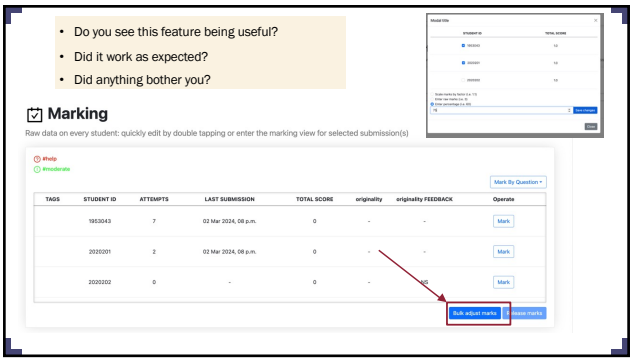

- Do you see this feature being useful?
- Did it work as expected?
- Did anything bother you?

15

## Task 4 –bulk adjust grades, dashboard

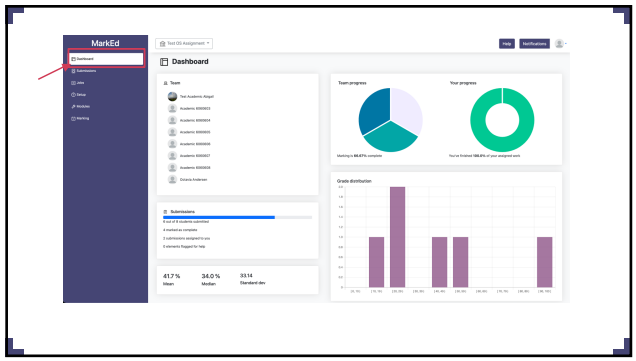
1. Go back to the "marking" tab.
2. Bulk adjust marks for the first three students.
   1. *Give these students 65%*
3. Bulk adjust marks for the last three students.
   1. *Give them a raw mark of 0.5*
4. View the dashboard for the assignment.
   1. *Take a look at the charts and statistics.*

16

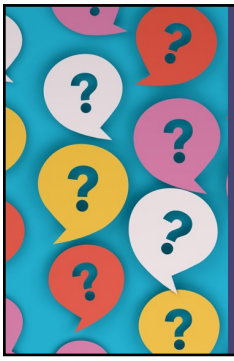



- Do you see this feature being useful?
- Did it work as expected?
- Did anything bother you?

17




18



## General comments

- What did you like the most?
- What did you like the least?
- What could be improved, and how?
- Any other suggestions?

19

## Questionnaire

On usability and potential impact

- LINK: https://forms.office.com/Pages/ResponsePage.aspx?id=sAafLmkWiUWHiRCgaTTcYcWwQnjee91KkTqFM_90YpxURFRDMkhCSVhMS0VCRFNHN1daMkpWWkVCNS4u

20

## THANK YOU

21

# Initial codes for thematic analysis

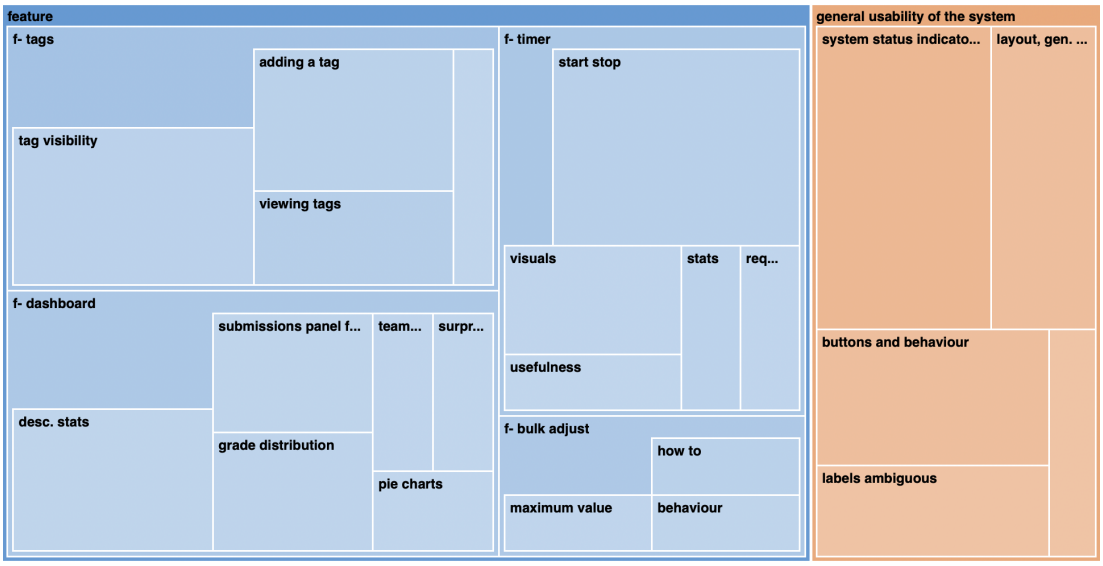



Figure C.1: Initial codes for the thematic analysis with box size indicating relative number of occurrences of a code

# Think aloud protocol with markers

## Plan

- Hi, thanks for coming!
- **PIS FORM For recording**
- **RECORD**
- *Intro:*
    - MarkEd is a tool for marking, moderation and feedback, similar to platforms used currently in the School of Informatics such as Blackboard Learn or Gradescope.
    - Today you will be evaluating some of the features offered by the tool as part of the marking process.
    - You will be given 5(12?) tasks to carry out following the Think Aloud protocol, followed by a questionnaire to collect some quantitative data based and some more open interview questions at the end.

  (Assist until the marker "home" screen.)
  Login details:
- *Think Aloud Protocol:*
    - I will give you a task to carry out on the interface. You will be asked to share your screen such that I can follow your train of thought. While you attempt to complete the task you are asked to "Tell me what you are thinking about as you work."

"In this observation, we are interested in what you think about as you perform the tasks we are asking you to do. In order to do this, I am going to ask you to talk aloud as you work on the task. What I mean by "talk aloud" is that I want you to tell me everything you are thinking from the first time you see the statement of the task until you finish the task. I would like you to talk aloud constantly from the time I give you the task until you have completed it. I don't want you to try to plan out what you say or try to explain to me what you are saying. Just act as if you were alone, speaking to yourself. It is most important that you keep talking. If you are silent for any long period of time, I will ask you to talk. Do you understand what I want you to do?"

    - You can share any reactions you have such as what your approach is to complete the task, what surprised you or what is confusing you. You are encouraged to ask questions aloud such as "Why is this button placed here?" but I will not be able to answer these at this time. The idea is to record what you are thinking with as little input from me as possible.
    - I will only observe and remind you to keep talking if you are quiet for some time.
    - This will take between 10 and 15 minutes.
- Tasks:
    - Practice task: login -> until marker interface ? (502 , mark502 ; 510 , mark510)
        - Log in using the credentials

  0. Select the CW 1 for the course "Software Engineering and Professional Practice".
  1. Go to the "marking" page.
  2. Start marking by question and select shuffle.
  3. Find the first un-marked submission.

4. Mark a submission by giving a mark, some feedback notes and save this as a draft.
5. Add a custom tag only visible to you. Call it “re-check diagram”.
6. Add a custom tag visible to everyone and name it “potentially AI generated”
7. Delete the first tag you created.
8. Select “complete” to finish marking this submission.
9. Stop the timer.
10. Examime the statistics generated.
11. Go to the dashboard.
12. Look at every white box element in turn and describe your thoughts.

| | Specific Qs | |
|---|---|---|
| Marking by question for coursework | - How often do you “mark by question” or “mark horizontally”?<br>- Is there a difference in your work pattern when marking by question? | Personal notes (prompts if the user is silent or struggling):<br>• *Tell me what you are thinking*<br>• *Please keep talking*<br>• *Please update me on what you are thinking*<br><br>• *Due to time constraints I’ll ask you to complete the next task …* |
| Tag | - For what purpose would you use tags as they appear now?<br>- Is there anything missing from the tag functionality as they appear now? | |
| Shuffle | - Can you think of a scenario in which you would want the shuffle functionality? | |
| Timer | - How useful do you think having an in-app timer is as opposed to using a simple clock or phone? | |
| Dashboard | - Any other stats you’d like to see on the dashboard? | |

- Final QUESTIONNAIRE:
- *Interview (ask additional questions if they come up)*.
- **THANK YOU**

## Teaching staff questionnaire

**Please note:** Questions 3 and 10 were omitted in the marker questionnaire.

# MarkEd: the tool for marking, moderation and feedback - MARKER and ACADEMIC questionnaire

Questionnaire to help evaluate work done on the marker interfaces for the tool.

* Required

## Introduction

1

***Please take a few minutes to read the following participant information sheet:***
https://uoe-my.sharepoint.com/:b:/g/personal/s1953043_ed_ac_uk/ESLyrKhLE2lDiqmaDPg59GsBdMFluCszQcoxj50mbdDddw

Please select "yes" if you agree to **ALL** of the statements below:

- I have read and understood the Participant Information Sheet for the above study, that I have had the opportunity to ask questions, and that any questions I had were answered to my satisfaction.
- My participation is voluntary, and that I can withdraw at any time without giving a reason. Withdrawing will not affect any of my rights.
- I consent to my anonymised data being used in academic publications and presentations.
- I understand that my anonymised data will be stored for the duration outlined in the Participant Information Sheet.
- I allow my data to be used in future ethically approved research.
- I agree to take part in this study.

*

- ◯ Yes
- ◯ No

2

Are you a Teaching Support Provider (Marker / Teaching Assistant) or senior academic (e.g. course organiser)? *

- ◯ Teaching Support Provider
- ◯ Academic
- ◯ Other

3

How long have you been involved in course organisation duties?

- ◯ 6 months or less
- ◯ 1 year or less
- ◯ 1-2 years
- ◯ 3-5 years
- ◯ over 5 years

4

How long have you been involved in marking duties?

- ◯ 6 months or less
- ◯ 7-12 months
- ◯ 1 year
- ◯ 2-3 years
- ◯ 4 years or more

## Feature specific questions

5

How helpful do you think the "**tag**" feature would be in the following areas? *

| | Very unhelpful | Somewhat unhelpful | Neither | Somewhat helpful | Very helpful |
|---|---|---|---|---|---|
| Saving time | ◯ | ◯ | ◯ | ◯ | ◯ |
| Ensuring fairness | ◯ | ◯ | ◯ | ◯ | ◯ |
| Ensuring high quality feedback | ◯ | ◯ | ◯ | ◯ | ◯ |

6

How helpful do you think the "**shuffle**" feature would be in the following areas? *

| | Very unhelpful | Somewhat unhelpful | Neither | Somewhat helpful | Very helpful |
|---|---|---|---|---|---|
| Saving time | ◯ | ◯ | ◯ | ◯ | ◯ |
| Ensuring fairness | ◯ | ◯ | ◯ | ◯ | ◯ |
| Ensuring high quality feedback | ◯ | ◯ | ◯ | ◯ | ◯ |

7

How helpful do you think the "**timer**" feature would be in the following areas? *

| | Very unhelpful | Somewhat unhelpful | Neither | Somewhat helpful | Very helpful |
|---|---|---|---|---|---|
| Saving time | ◯ | ◯ | ◯ | ◯ | ◯ |
| Ensuring fairness | ◯ | ◯ | ◯ | ◯ | ◯ |
| Ensuring high quality feedback | ◯ | ◯ | ◯ | ◯ | ◯ |

# System Usability Scale (SUS)

A standard questionnaire used to assess the usability of a product.

8

If you were given the choice as to what system to use for your assignments, please indicate your opinion regarding these 10 statements. *

| | STRONGLY disagree | disagree | neutral | agree | STRONGLY agree |
|---|---|---|---|---|---|
| I think that I would like to use this system (MarkEd) frequently. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I found the system unnecessarily complex. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I thought the system was easy to use. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I think that I would need the support of a technical person to be able to use this system. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I found the various functions in this system were well integrated. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I thought there was too much inconsistency in this system. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I would imagine that most students would learn to use this system very quickly. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I found the system very cumbersome to use. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I felt very confident using the system. | ◯ | ◯ | ◯ | ◯ | ◯ |
| I needed to learn a lot of things before I could get going with this system. | ◯ | ◯ | ◯ | ◯ | ◯ |

## Potential Impact for the School of Informatics

Looking at the potential benefits of MarkEd for the School of Informatics

9

**Overall,**
How helpful do you think the tool would be in the following areas in **your own marking**?

| | Very unhelpful | Somewhat unhelpful | Neither | Very helpful | Somewhat helpful |
|---|---|---|---|---|---|
| Saving time | ◯ | ◯ | ◯ | ◯ | ◯ |
| Ensuring fairness | ◯ | ◯ | ◯ | ◯ | ◯ |
| Error-prevention (avoiding mis-clicks) | ◯ | ◯ | ◯ | ◯ | ◯ |
| Combining strengths of other marking tools / platforms | ◯ | ◯ | ◯ | ◯ | ◯ |
| Understanding your own progress in marking | ◯ | ◯ | ◯ | ◯ | ◯ |
| Leaving “to-do” notes for future work | ◯ | ◯ | ◯ | ◯ | ◯ |
| Ensuring high quality feedback | ◯ | ◯ | ◯ | ◯ | ◯ |
| Asking for assistance from (co-)lecturers or TAs. | ◯ | ◯ | ◯ | ◯ | ◯ |

**[FOR ACADEMICS ONLY]**
**Overall,**
How helpful do you think the tool would be for the **markers you coordinate**, in the following areas?

| | Very unhelpful | Somewhat unhelpful | Neither | Very helpful | Somewhat helpful |
|---|---|---|---|---|---|
| Saving time | ◯ | ◯ | ◯ | ◯ | ◯ |
| Ensuring fairness | ◯ | ◯ | ◯ | ◯ | ◯ |
| Error-prevention (avoiding mis-clicks) | ◯ | ◯ | ◯ | ◯ | ◯ |
| Combining strengths of other marking tools / platforms | ◯ | ◯ | ◯ | ◯ | ◯ |
| Understanding your own progress in marking | ◯ | ◯ | ◯ | ◯ | ◯ |
| Leaving "to-do" notes for future work | ◯ | ◯ | ◯ | ◯ | ◯ |
| Ensuring high quality feedback | ◯ | ◯ | ◯ | ◯ | ◯ |
| Asking for assistance from (co-)lecturers or TAs. | ◯ | ◯ | ◯ | ◯ | ◯ |

## Impact scores for specific features

| Feature | Area | Mean | Median | Std. Dev. | Min | Max |
|---|---|---|---|---|---|---|
| Custom tag | Saving time | 4.0 | 4 | 0.71 | 3 | 5 |
| | Ensuring fairness | 3.8 | 4 | 0.84 | 3 | 5 |
| | **Ensuring high quality feedback** | 4.0 | 4 | 0.71 | 3 | 5 |
| Shuffle | Saving time | 3.0 | 3 | 0.00 | 3 | 3 |
| | **Ensuring fairness** | 4.0 | 4 | 1.00 | 3 | 5 |
| | Ensuring high quality feedback | 3.0 | 3 | 0.00 | 3 | 3 |
| Timer | **Saving time** | 3.6 | 4 | 1.14 | 2 | 5 |
| | Ensuring fairness | 2.6 | 3 | 1.14 | 1 | 4 |
| | Ensuring high quality feedback | 3.0 | 3 | 0.71 | 2 | 4 |

Table C.3: Potential impact of specific features on various areas of marking with the area the feature targeted being in **bold** font

## SUS questionnaire full scores

| Participant | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 | **Score** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A1 | 2 | 1 | 2 | 3 | 1 | 2 | 3 | 3 | 2 | 3 | 55 |
| A2 | 3 | 3 | 3 | 4 | 3 | 3 | 3 | 3 | 3 | 4 | 80 |
| M1 | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 80 |
| M2 | 3 | 3 | 4 | 3 | 4 | 4 | 4 | 3 | 2 | 3 | 82.5 |
| M3 | 3 | 3 | 4 | 4 | 3 | 3 | 3 | 4 | 4 | 3 | 85 |
| S1 | 3 | 4 | 4 | 4 | 3 | 4 | 4 | 4 | 4 | 4 | 95 |
| S2 | 4 | 3 | 3 | 4 | 4 | 3 | 4 | 4 | 4 | 4 | 92.5 |
| S3 | 4 | 3 | 3 | 4 | 4 | 4 | 2 | 4 | 3 | 4 | 87.5 |
| S4 | 3 | 4 | 4 | 4 | 3 | 3 | 4 | 3 | 1 | 3 | 80 |
| S5 | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 77.5 |
| S6 | 3 | 3 | 3 | 4 | 4 | 3 | 3 | 1 | 3 | 4 | 77.5 |
| S7 | 3 | 3 | 3 | 4 | 3 | 4 | 3 | 3 | 3 | 4 | 82.5 |
| S8 | 3 | 4 | 3 | 4 | 0 | 4 | 4 | 4 | 4 | 4 | 85 |
| S9 | 2 | 3 | 3 | 4 | 3 | 3 | 3 | 2 | 2 | 4 | 72.5 |
| S10 | 4 | 4 | 3 | 4 | 3 | 3 | 4 | 4 | 4 | 3 | 90 |
| S11 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 75 |
| S13 | 2 | 3 | 2 | 3 | 2 | 3 | 0 | 3 | 0 | 3 | 52.5 |
| S15 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 97.5 |
| S16 | 2 | 3 | 3 | 4 | 3 | 2 | 3 | 3 | 3 | 4 | 75 |
| S17 | 4 | 0 | 4 | 4 | 4 | 4 | 4 | 0 | 4 | 4 | 80 |
| S18 | 2 | 3 | 3 | 4 | 3 | 2 | 3 | 3 | 3 | 4 | 75 |
| S19 | 3 | 4 | 3 | 4 | 3 | 2 | 3 | 3 | 3 | 4 | 80 |
| S20 | 2 | 4 | 2 | 4 | 2 | 3 | 2 | 4 | 2 | 4 | 72.5 |
| S22 | 4 | 4 | 4 | 4 | 3 | 4 | 3 | 3 | 4 | 3 | 90 |
| S23 | 3 | 3 | 3 | 4 | 3 | 2 | 3 | 1 | 3 | 3 | 70 |
| S24 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 100 |
| **Mean** | 3.1 | 3.2 | 3.2 | 3.8 | 3.0 | 3.2 | 3.2 | 3.0 | 3.0 | 3.6 | |
| **Median** | 3 | 3 | 3 | 4 | 3 | 3 | 3 | 3 | 3 | 4 | |

Table C.4: Complete participant SUS Scores

# Appendix D

# Full requirements tables

## Student pages requirements full table

| # | Diss | Page | OG Req. | Description | Impact | Confidence | Ease | ICE score |
|---|---|---|---|---|---|---|---|---|
| 1.1 | XW | 35 | 1.3 | Student should be able to view their submitted files. | 3.5 | 3 | 4 | 42.0 |
| 1.2 | AG | 46 | 4.4 | Student should receive confirmation email when assignment is submitted. | 2 | 3 | 4 | 24.0 |
| 1.3 | XW | 35 | | Assignments with no submissions should be at the top. | 1 | 4 | 4 | 16.0 |
| 1.4 | XW | 35 | | 12 o'clock should not be shown as noon. | 1 | 4 | 4 | 16.0 |
| 1.5 | XW | 35 | | Instruction for the assignment to be submitted should be visible. | 1 | 4 | 4 | 16.0 |
| 1.6 | XW | 35 | | 'Back' button is needed on interior pages. | 1 | 4 | 4 | 16.0 |
| 1.7 | XW | 35 | | An explanation of sandwich structure should be provided. | 1 | 4 | 4 | 16.0 |
| 1.8 | AG | 40 | 4.4 | Student notifications for marks and feedback should be added. | 2 | 2 | 3 | 12.0 |
| 1.9 | AG | 46 | 1.9 | Coursework result calculator should be added. | 2 | 2 | 1 | 4.0 |

# Marker pages requirements full table

| # | Diss | Page | OG Req. | Description | Impact | Confidence | Ease | ICE score |
|---|---|---|---|---|---|---|---|---|
| 2.1 | XS | 31 | | The ”Jobs” page function should be clearer. | 4 | 4 | 4 | 64.0 |
| 2.2 | CS | 27 | | Shuffle questions should be completed. | 4 | 3 | 4 | 48.0 |
| 2.3 | AG | 49 | 2.4 | Dashboard of marking progress should be completed. | 4 | 4 | 3 | 48.0 |
| 2.4 | AG | 76 | | It should be possible to adjust the space taken up by a submission on the marking page. | 4 | 3 | 4 | 48.0 |
| 2.5 | AG | 95 | | ”Days Late” column should be added to the marking data column. | 3 | 4 | 4 | 48.0 |
| 2.6 | CS | 28 | | Timed marking should be supported. | 4 | 3 | 3 | 36.0 |
| 2.7 | AG | 49 | 2.7 | Analytics on student performance should be added. | 2 | 4 | 4 | 32.0 |
| 2.8 | CS | 35 | | The system should handle ”no submission”. | 2 | 4 | 4 | 32.0 |
| 2.9 | XW | 34 | 2.8 | The user should be able to create custom tags. | 3.5 | 3 | 3 | 31.5 |
| 2.10 | AG | 49 | 3.2 | Ability to modify marks across multiple submissions at once should be added. | 2.5 | 4 | 3 | 30.0 |
| 2.11 | CS | 27 | 2.9 | Review marks should be completed. | 3 | 3 | 3 | 27.0 |
| 2.12 | XW (+CS27) | 34 | | Filtering submissions by tag or similar score should be added. | 2 | 3 | 4 | 24.0 |
| 2.13 | AG | 92 | 2.3 | Splitting submissions across markers: per submission / per question / custom should be completed. | 4 | 3 | 2 | 24.0 |
| 2.14 | XS | 29 | | Guidance for using the tool should be available. | 2 | 4 | 3 | 24.0 |

| Marker pages requirements continued | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| # | Diss | Page | OG Req. | Description | Impact | Confidence | Ease | ICE score |
| 2.15 | XS | 31 | 2.3 | Assigning work needs to be clearer. | 3.5 | 3 | 2 | 21.0 |
| 2.16 | AG | 36 | 2.12 | Custom user role support should be added. | 2 | 3 | 3 | 18.0 |
| 2.17 | XS | 32 | 2.10 | Conflicting permission setting between groups and individual markers should be prevented. | 2 | 3 | 3 | 18.0 |
| 2.18 | AG | 36 | 2.10 | Custom permissions for read / write access to submissions should be completed. | 2 | 3 | 3 | 18.0 |
| 2.19 | XS | 31 | | Emailing functionality should be completed. | 2 | 3 | 3 | 18.0 |
| 2.20 | CS | 28 | | Anchor and compare assignments should be completed. | 4 | 2 | 2 | 16.0 |
| 2.21 | XW (+HA34) | 34 | | The feedback bank should be added. | 4 | 2 | 2 | 16.0 |
| 2.22 | AG | 93 | 4.2 | Easy export of student data in a custom format should be added. | 4 | 2 | 2 | 16.0 |
| 2.23 | XW | 34 | | Back button should be added to ”Compare” page. | 1 | 4 | 4 | 16.0 |
| 2.24 | XW | 34 | | Message should notify the user that their work was saved. | 1 | 4 | 4 | 16.0 |
| 2.25 | XW | 34 | | Hover hints should be used for tag explanation. | 1 | 4 | 4 | 16.0 |
| 2.26 | XW | 34 | | The ”Compare” page should have a button to access the assignments displayed. | 1 | 4 | 4 | 16.0 |
| 2.27 | XS | 28 | | Submission enable/disable button should be visible and provide feedback to the user. | 1 | 4 | 4 | 16.0 |
| 2.28 | AG | 72 | | Next/previous submission buttons should be clearer. | 1 | 4 | 4 | 16.0 |
| 2.29 | AG | 95 | | ”Marking Data” should be higher up in the menu bar. | 1 | 4 | 4 | 16.0 |
| 2.30 | CS | 27 | 2.9 | Double marking should be added. | 2.5 | 3 | 2 | 15.0 |
| 2.31 | AG | | 1.5 | Anonymised submissions should be supported. | 1.5 | 3 | 3 | 13.5 |
| 2.32 | AG | 49 | 3.1 | Analytics on marker performance should be added. | 1.5 | 3 | 3 | 13.5 |
| 2.33 | AG | 46 | 1.6 | Group submission support should be added. | 2 | 3 | 2 | 12.0 |
| 2.34 | AG | 49 | 2.7 | Analytics on student performance should be added. | 2 | 2 | 3 | 12.0 |
| 2.35 | XW | 34 | | The marker should receive reminders to save work. | 1 | 3 | 4 | 12.0 |
| 2.36 | AG | 72 | | Autosave feature should be added. | 1 | 3 | 4 | 12.0 |

| Marker pages requirements continued | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| # | Diss | Page | OG Req. | Description | Impact | Confidence | Ease | ICE score |
| 2.37 | XW | 34 | 2.3 | Marker should only see the submissions assigned to them. | 2.5 | 2 | 2 | 10.0 |
| 2.38 | XW | 34 | | Alternatives to the sandwich structure should be added. | 1 | 3 | 3 | 9.0 |
| 2.39 | XW | 34 | 2.13 | Ability to perform similarity checks on submissions should be added. | 1.5 | 1 | 4 | 6.0 |
| 2.40 | CS | 35 | | Search for a submission should be added. | 1 | 3 | 2 | 6.0 |
| 2.41 | AG | 94 | | Automated workflows should be completed. | 3 | 2 | 1 | 6.0 |
| 2.42 | AG | 95 | | Support for keyboard shortcuts should be added. | 1 | 2 | 3 | 6.0 |
| 2.43 | AG | 46 | | Support for assignment deadlines with extensions should be added. | 1 | 3 | 2 | 6.0 |
| 2.44 | XW | 34 | | Support for holistic / criteria-based marking should be added. | 2 | 2 | 1 | 4.0 |
| 2.45 | AG | 93 | | Annotating submissions should be supported. | 1 | 2 | 2 | 4.0 |
| 2.46 | AG | 46 | 2.6 | Automarker support should be added. | 3 | 1 | 1 | 3.0 |
| 2.47 | AG | 49 | 4.4 | Regular backups of data should be added. | 1.5 | 1 | 2 | 3.0 |
| 2.48 | XS | 30 | | The system should support multiple simultaneous users. | 3 | 1 | 1 | 3.0 |
| 2.49 | AG | 49 | 2.11 | Sandboxing of student submissions should be added. | 2.5 | 1 | 1 | 2.5 |
| 2.50 | AG | 49 | 4.1 | Possibility of integrations with existing tools should be added. | 2.5 | 1 | 1 | 2.5 |
| 2.51 | AG | 49 | 4.3 | Generation of analysis for Board of Examiners meetings. | 1 | 1 | 2 | 2.0 |
| 2.52 | AG | 49 | 2.5 | Command-line interface access to the system should be added. | 1.5 | 1 | 1 | 1.5 |
| 2.53 | CS | 27 | | Marker chat should be added. | 1 | 1 | 1 | 1.0 |

# Appendix E

# My implementation screenshots

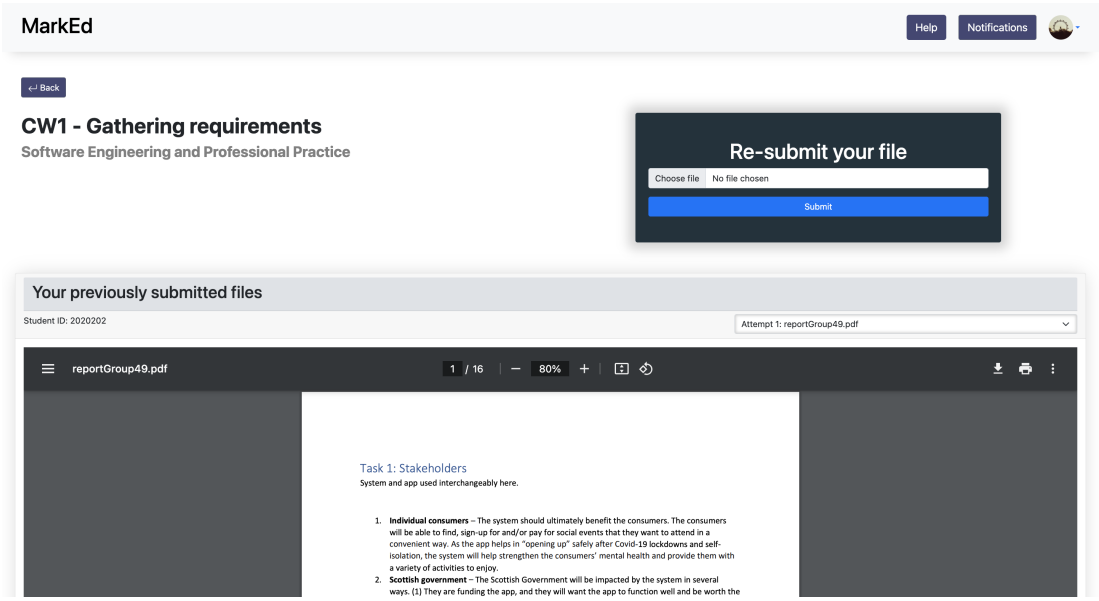



Figure E.1: Submit page with panel to view previous submissions and course and course work title

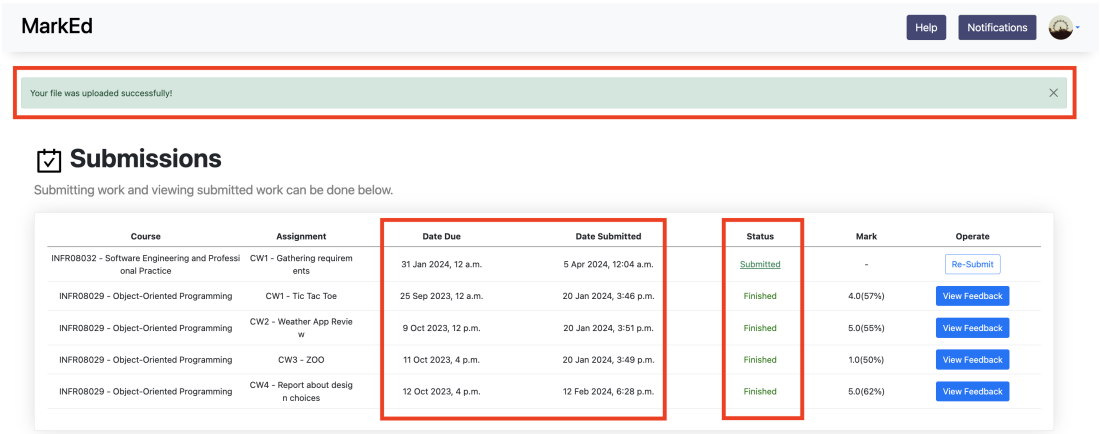



Figure E.2: Student Home page with submission confirmation

| Date Due | Date Submitted |
|---|---|
| 24 Sep 2023, 4 p.m. | - |
| 12 Oct 2023, 4 p.m. | - |
| 25 Sep 2023, 12 a.m. | 17 Oct 2023, 7:14 p.m. |
| 9 Oct 2023, 12 p.m. | 26 Nov 2023, 4:13 p.m. |

(a) British localisation

| Date Due | Date Submitted |
|---|---|
| 24. September 2023, 16:00 | - |
| 12. Oktober 2023, 16:00 | - |
| 25. September 2023, 00:00 | 17. Oktober 2023, 19:14 |
| 9. Oktober 2023, 12:00 | 26. November 2023, 16:13 |

(b) German localisation

Figure E.3: Comparison of localisation for the time formats on the student Home page

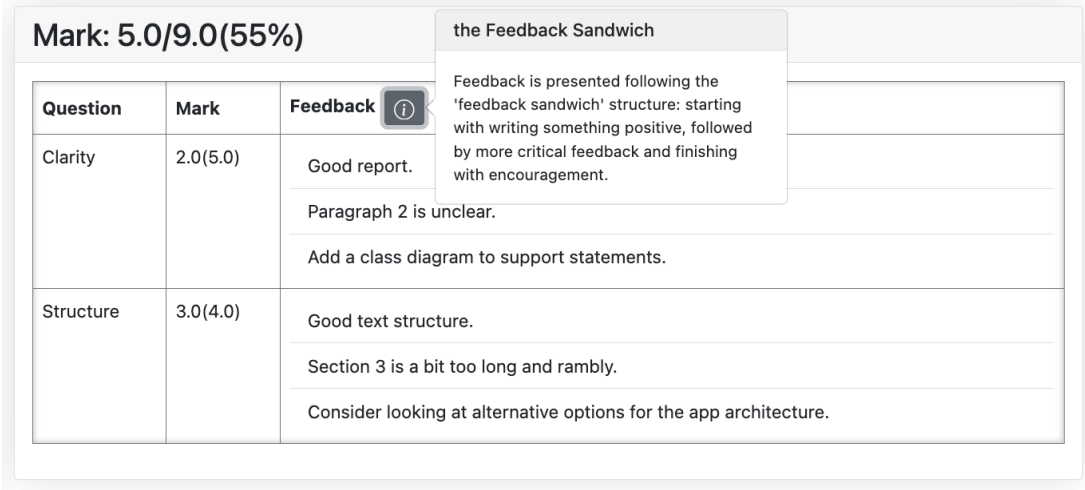



Figure E.4: The feedback sandwich explanation popover on the Feedback page

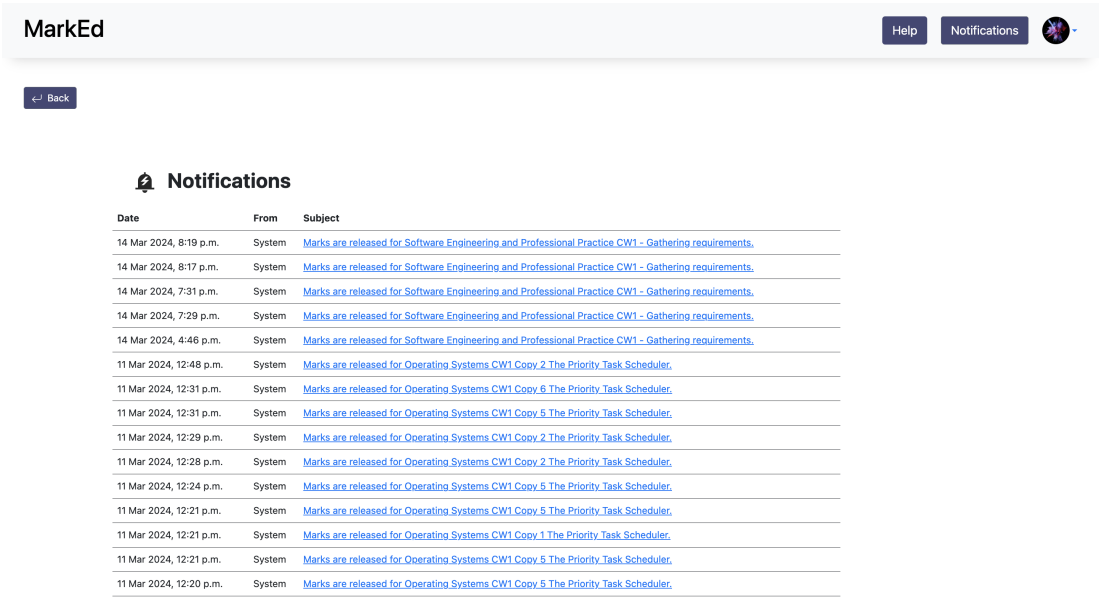



Figure E.5: Student Notification page

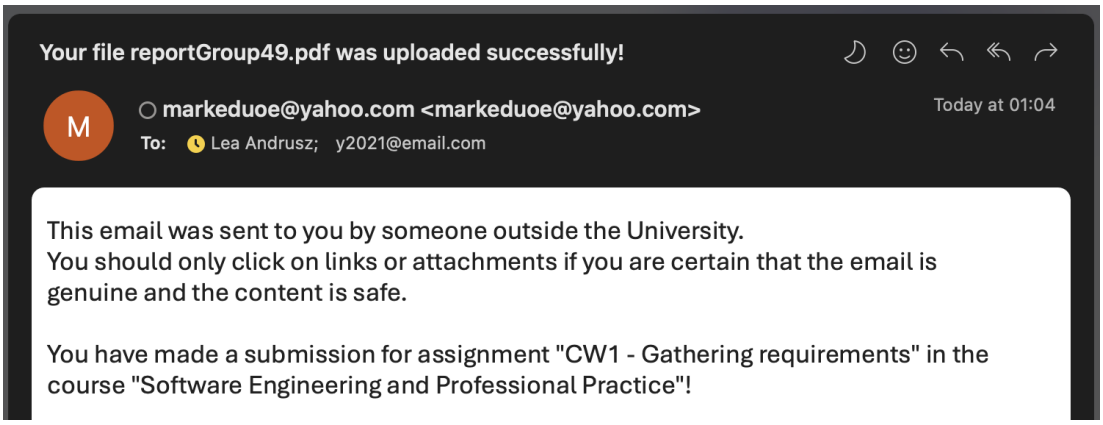



Figure E.6: Screenshot of automatic email notification of successful file submission

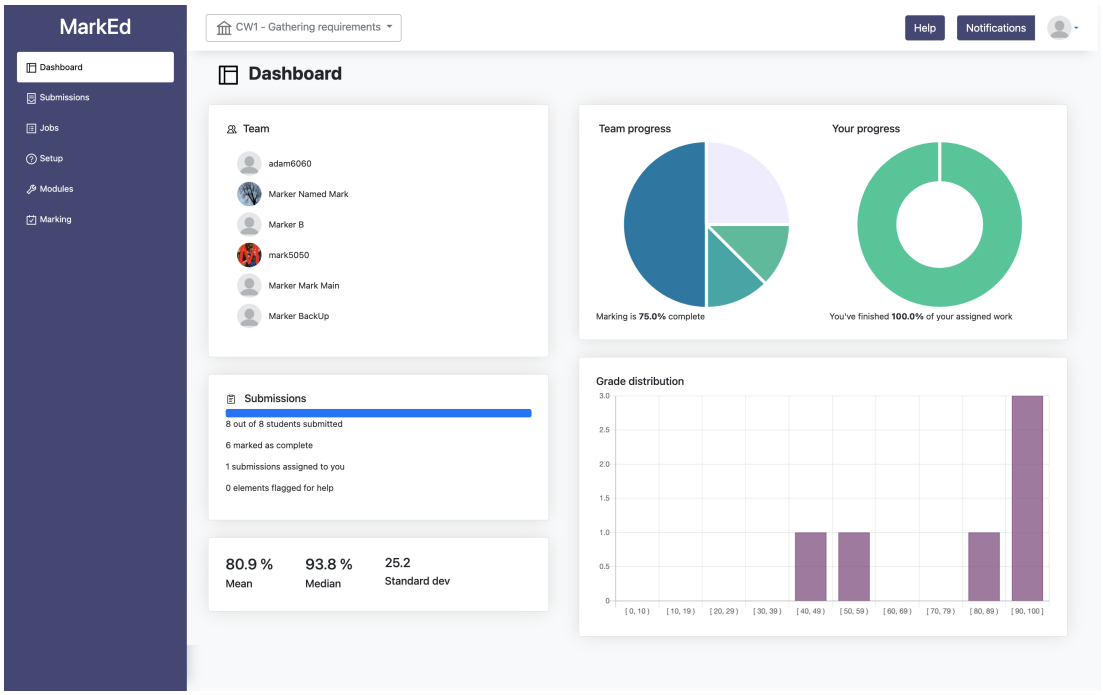



Figure E.7: Teaching staff dashboard with cohort performance analytics

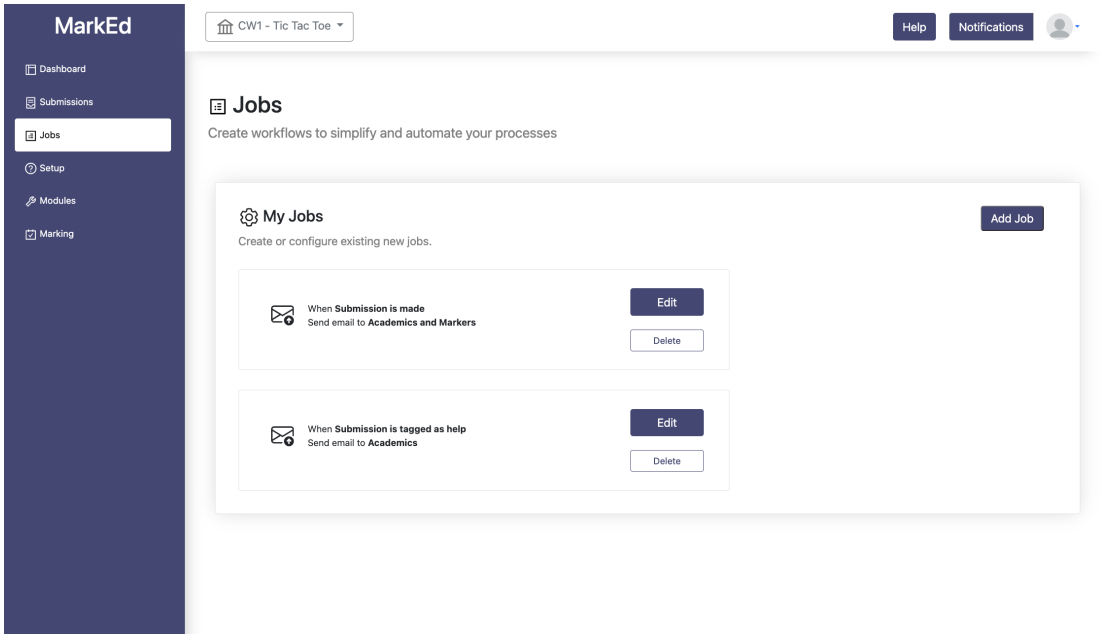



Figure E.8: Re-designed Jobs page

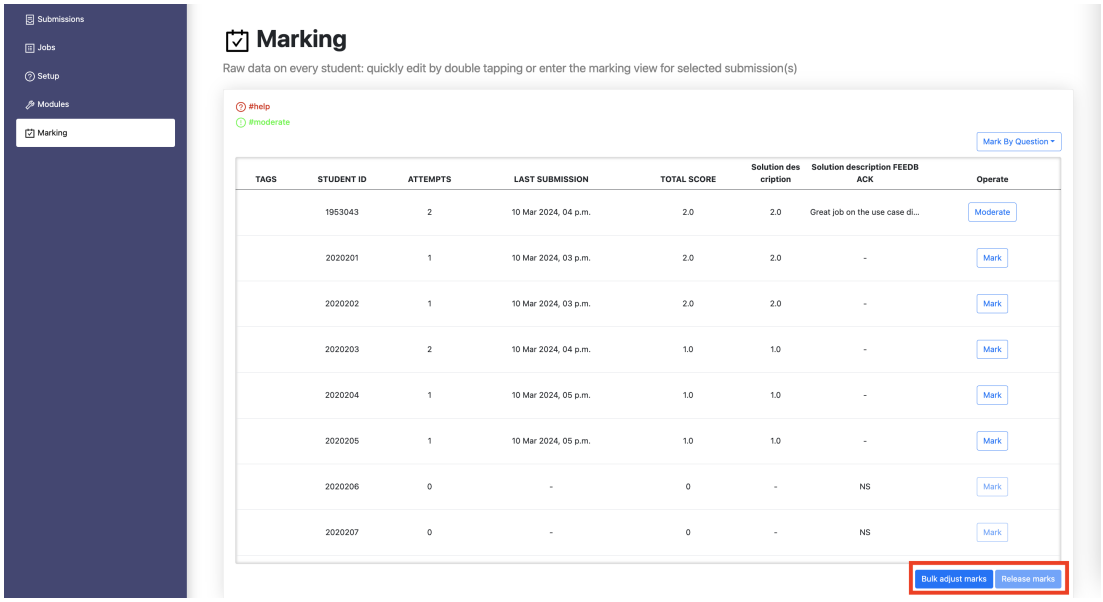



Figure E.9: Placement of the “bulk adjust marks” and “release marks” buttons

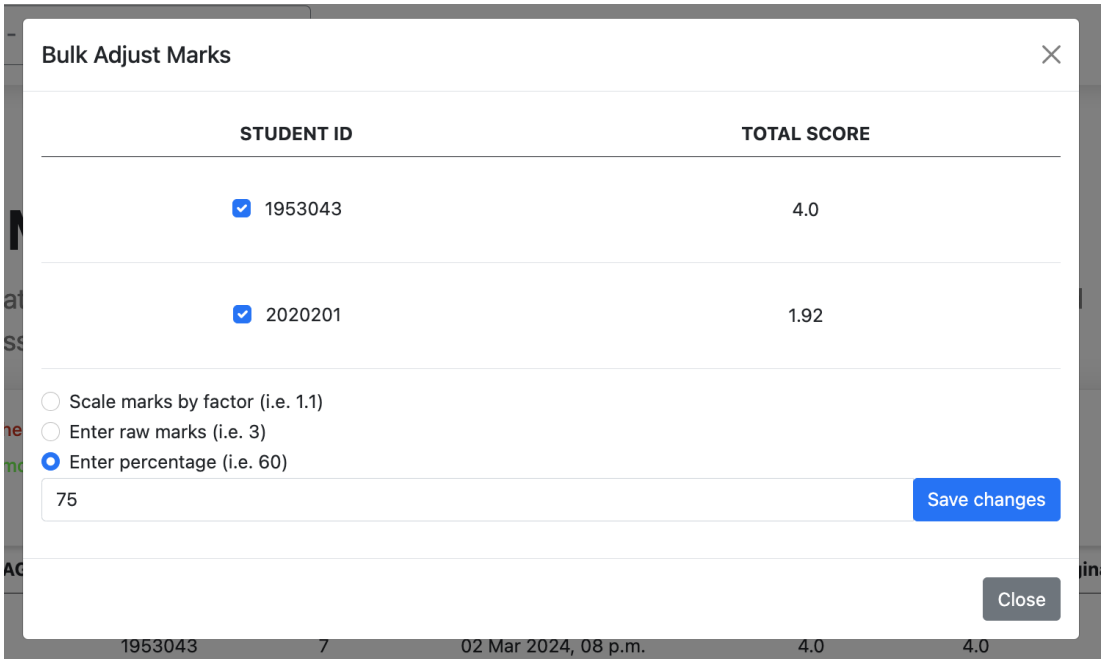



Figure E.10: Bulk adjust marks modal

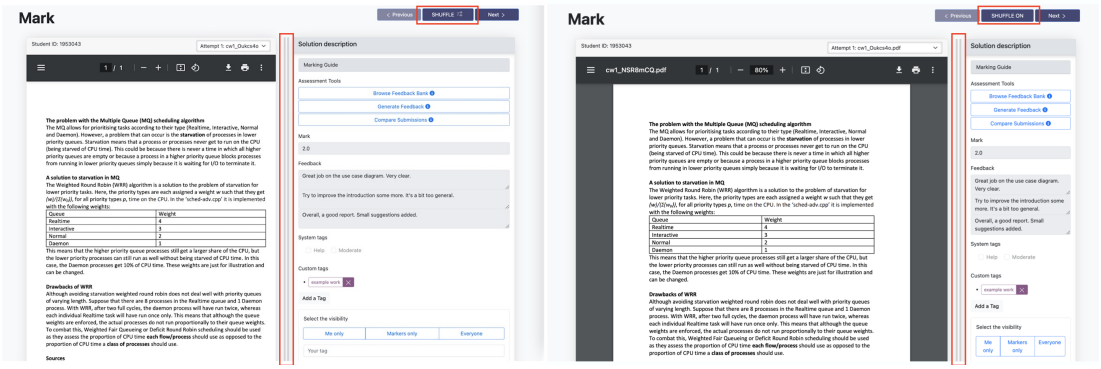



Figure E.11: Shuffle button toggled on/off and panel width adjusted using slider

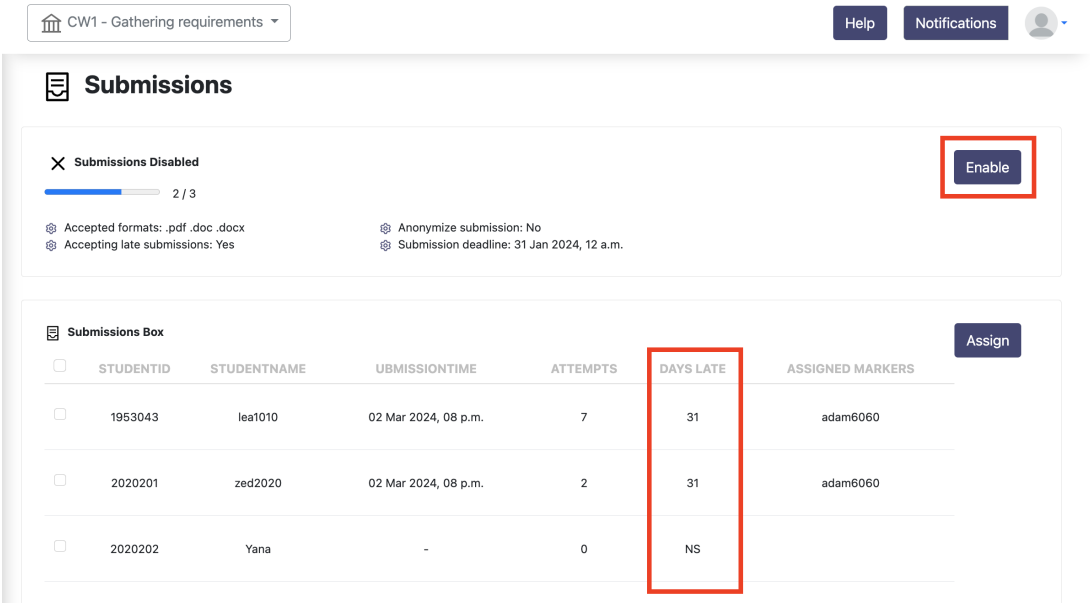



Figure E.12: Submissions page table with days late and no submissions

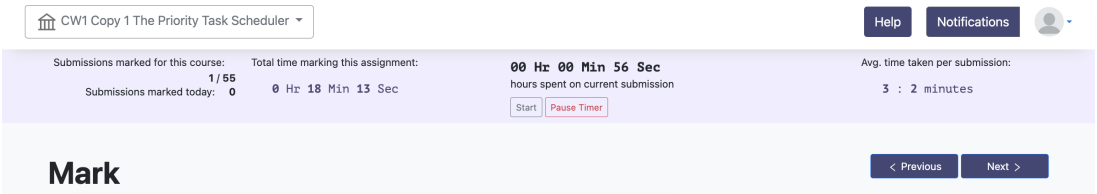



Figure E.13: Timer ribbon with the marker's personal statistics generated

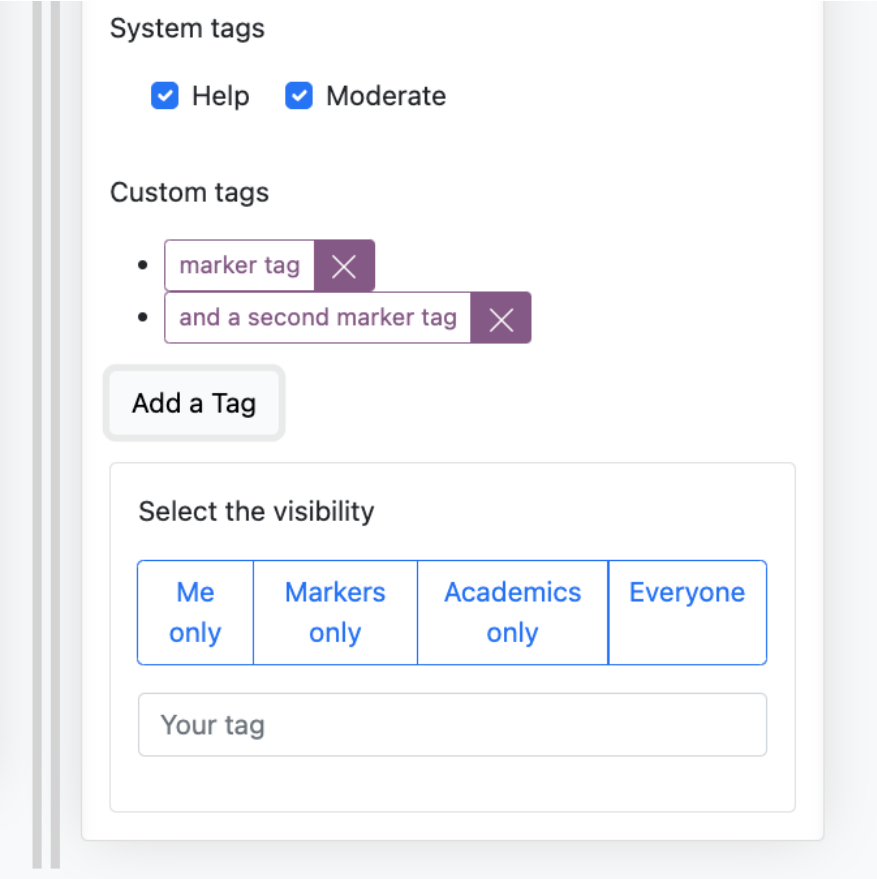



Figure E.14: Custom tags drop-down with visibility options

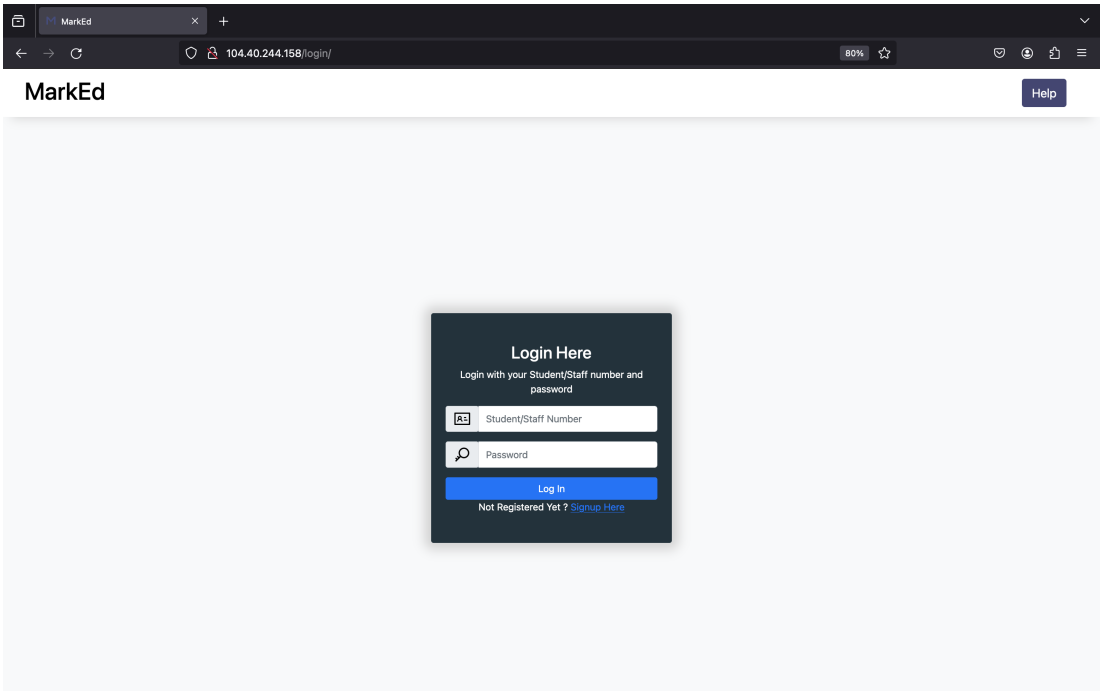



Figure E.15: Screenshot of Marked accessible from the browser

# Andrius original requirements

| Category | | Functional Requirement | Priority | Reason |
|---|---|---|---|---|
| **1. Submission** | 1.1 | Submission format enforcement (e.g. only PDF files accepted) | High | I |
| | 1.2 | Ability to submit | High | S |
| | 1.3 | Ability for students to view submissions | High | S |
| | 1.4 | Timestamp submissions, retain all submissions from the student | High | C |
| | 1.5 | Anonymized submissions | Low | NI |
| | 1.6 | Group submission support | Low | NL |
| | 1.7 | Ability for students to contact markers within the system to enquire about feedback | Low | NI |
| | 1.8 | Ability for students to view marks and feedback in the system | High | S |
| | 1.9 | Potential coursework result calculator | Very Low | L |
| **2. Marking** | 2.1 | Ability to reuse feedback comments quickly | Medium | SN |
| | 2.2 | Ability to record feedback | High | C |
| | 2.3 | Splitting submissions across markers: per submission / per question / custom | High | SI |
| | 2.4 | Dashboard of marking progress | High | S |
| | 2.5 | Command-line interface access to the system | Low | N |
| | 2.6 | Automarker support | Medium | SN |
| | 2.7 | Analytics on student performance | Low | SN |
| | 2.8 | Ability to custom flag submissions by markers for further support | High | SI |
| | 2.9 | Double marking / double blind marking | Low | LN |
| | 2.10 | Custom permissions for read / write access to submissions | Medium | C |
| | 2.11 | Sandboxing of student submissions | High | SC |
| | 2.12 | Custom user role support | Medium | LI |
| | 2.13 | Plagiarism checks | Very Low | NL |
| **3. Moderation** | 3.1 | Analytics on marker performance | Low | N |
| | 3.2 | Ability to modify feedback and marks across multiple submissions at once | Medium | SN |
| **4. Administrative** | 4.1 | Possibility of integrations with existing tools | High | SI |
| | 4.2 | Easy export of student data in a custom format | High | S |
| | 4.3 | Generation of analysis for Board of Examiners meetings | Very Low | LN |
| | 4.4 | Regular backups of data | Low | C |
| | 4.4 | Student notifications for marks and feedback / bulk emails | High | S |
| | 4.5 | Audit log functionality of system users | Low | LN |

Figure E.16: Andrius' original requirements

# Appendix F

# Andrius designs

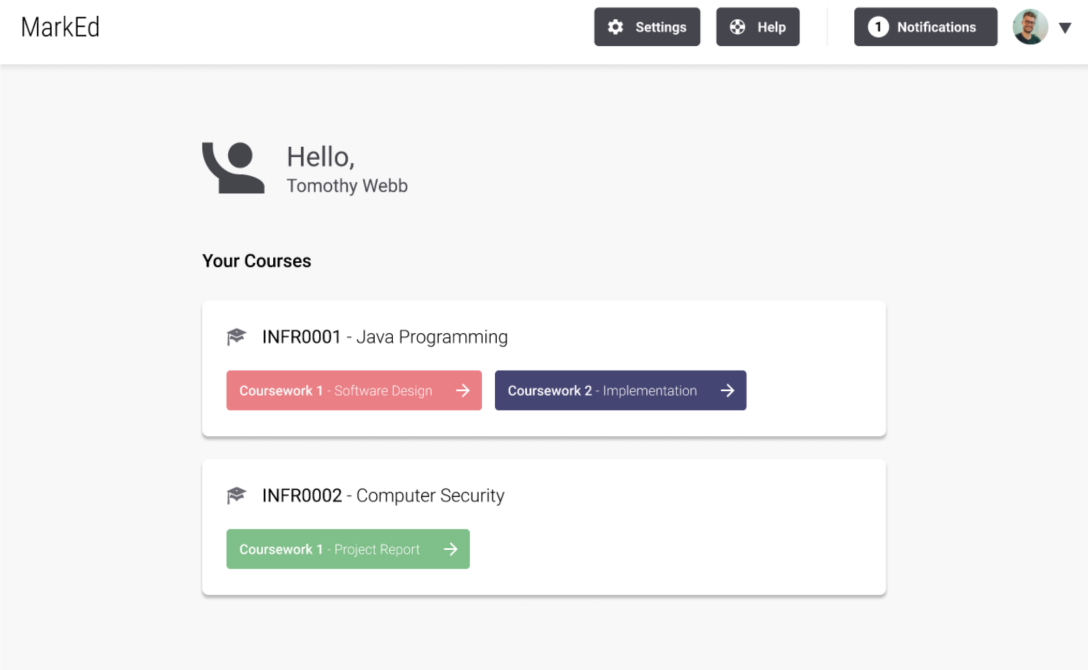



Figure F.1: Home page

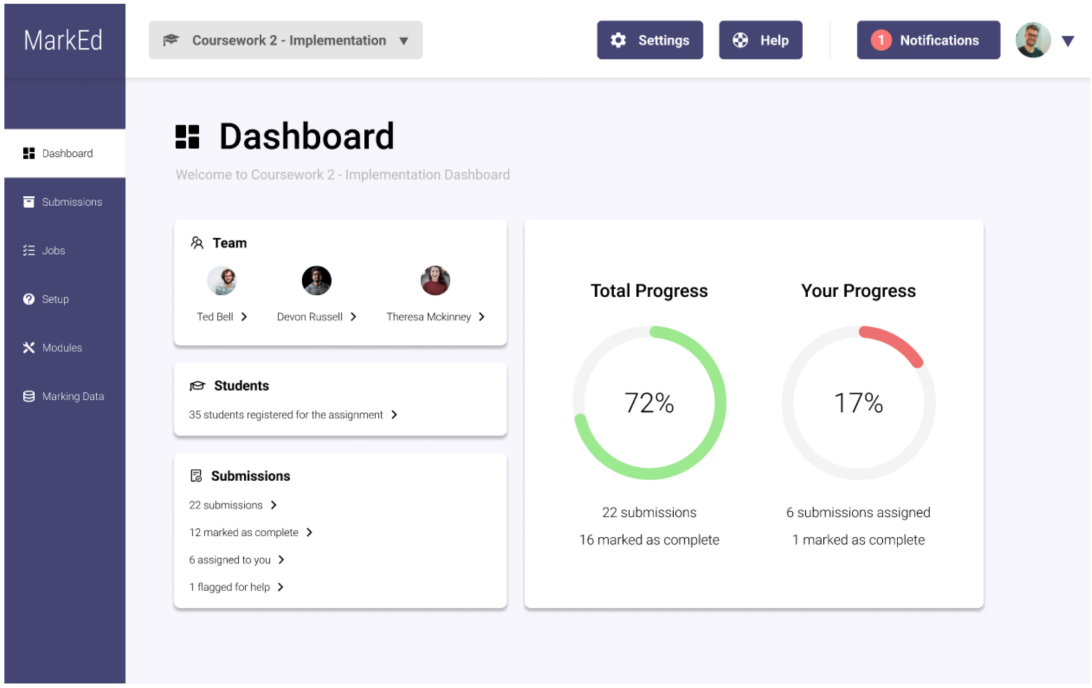



Figure F.2: Dashboard Page

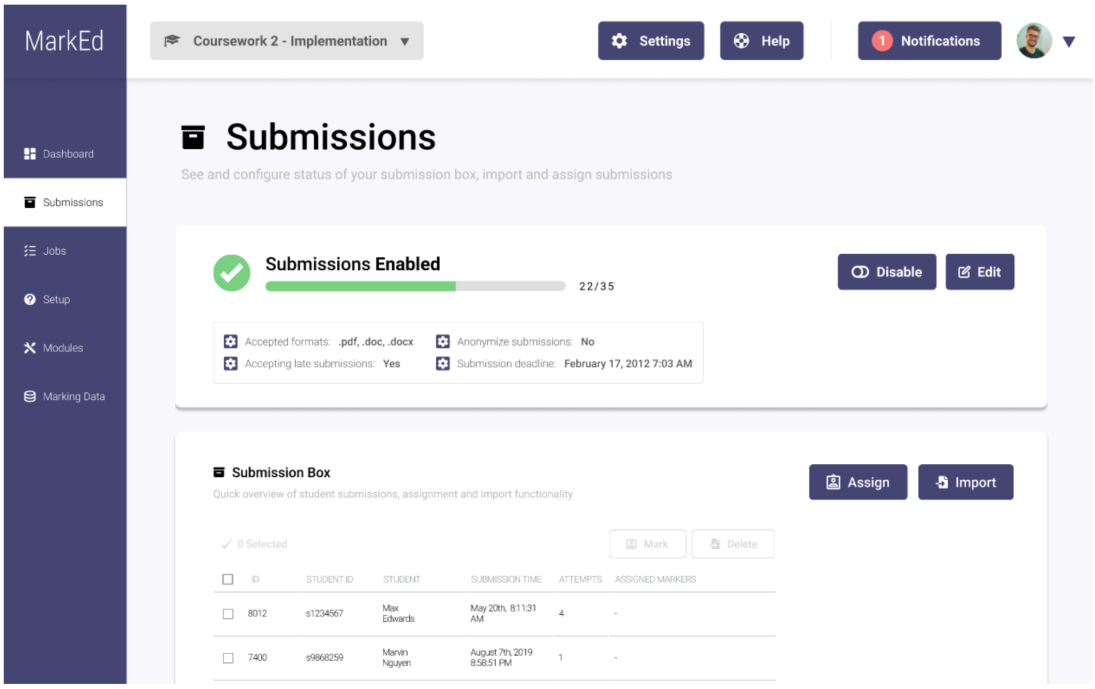



Figure F.3: Submissions Page

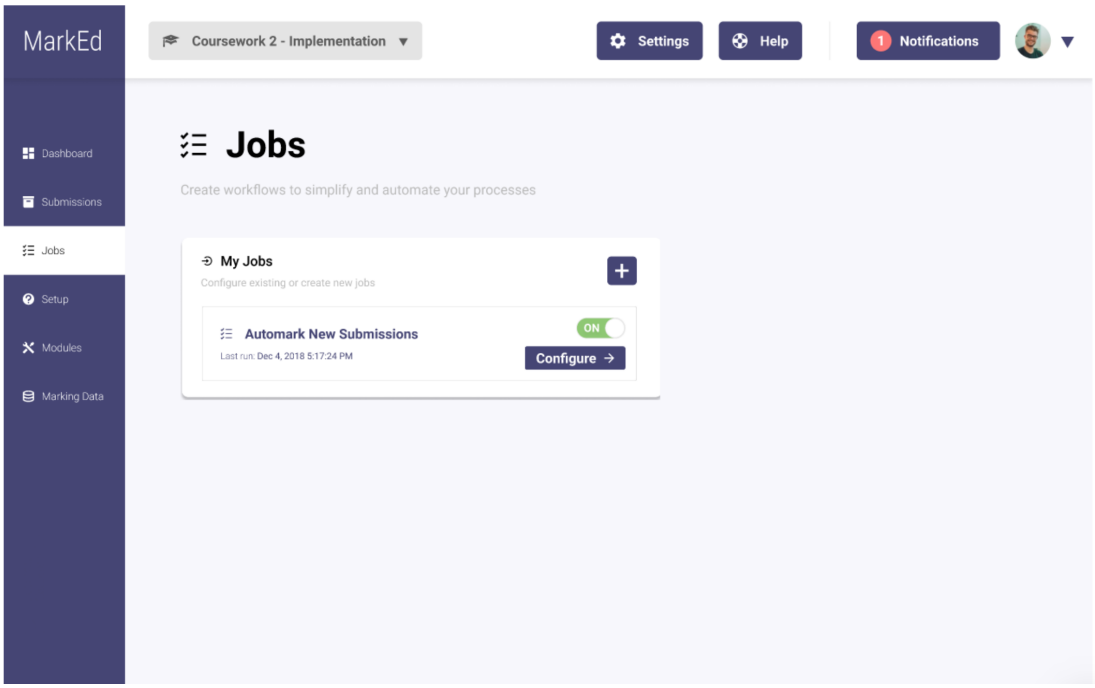



Figure F.4: Jobs Page

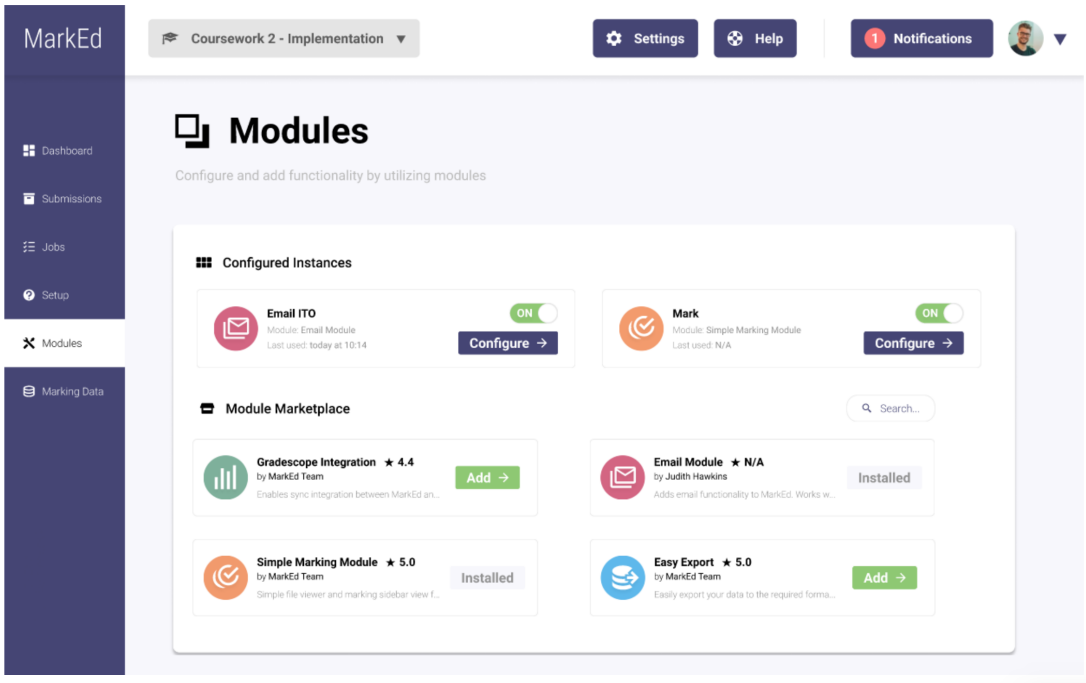



Figure F.5: Modules Page

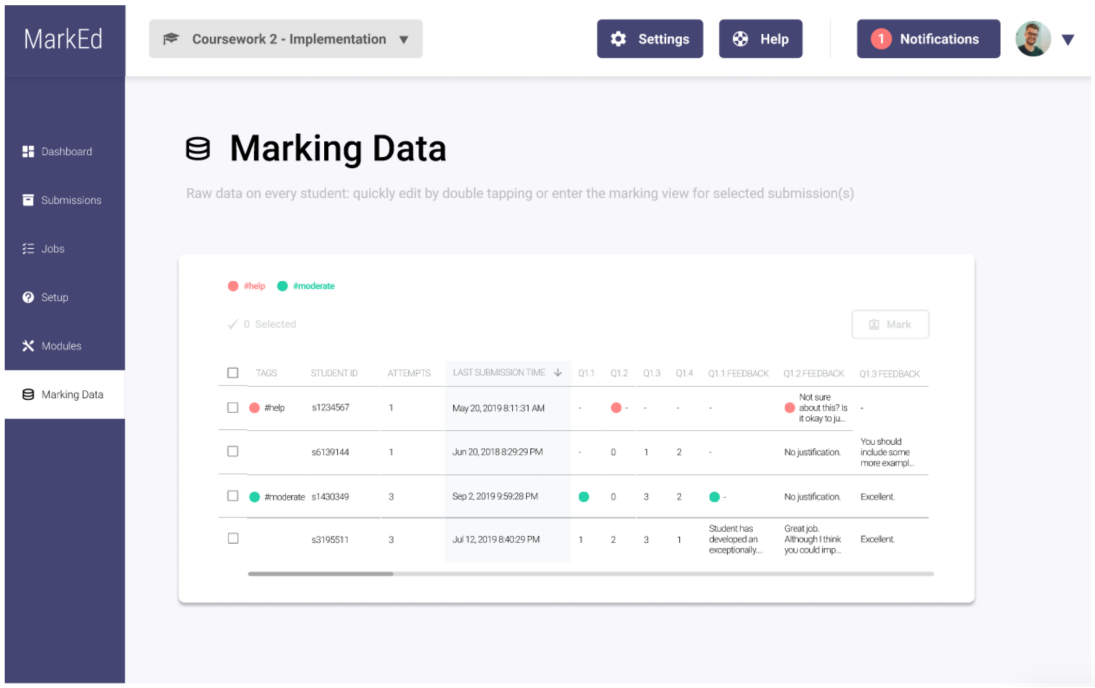



Figure F.6: Marking Data Page

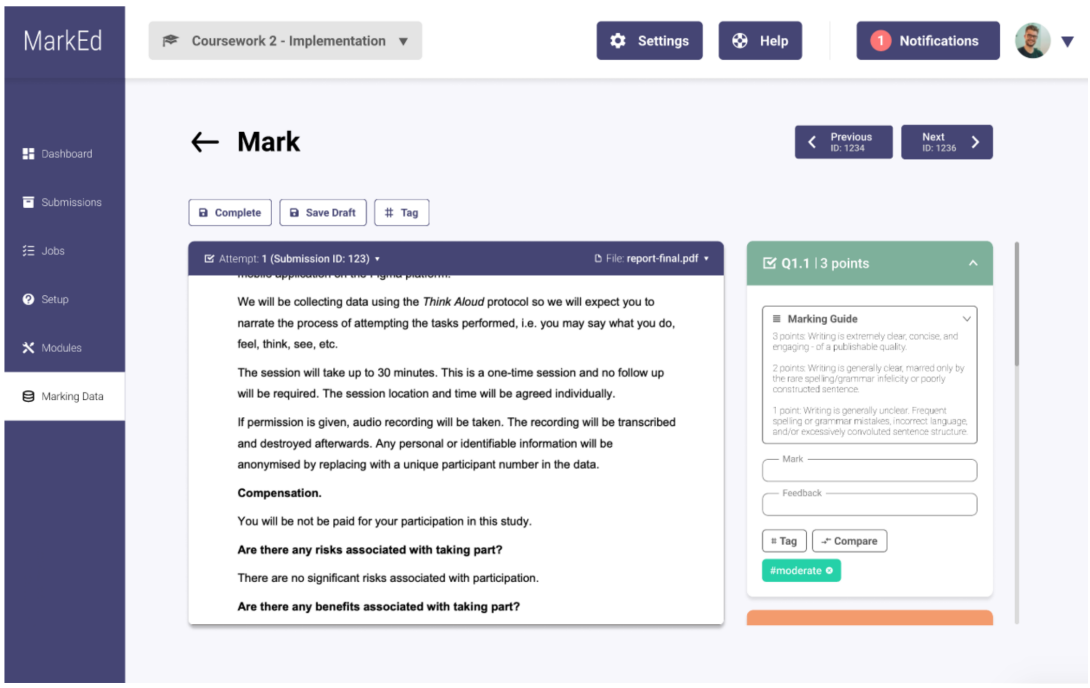



Figure F.7: Mark Page

# Chris' strategies for fair and efficient marking

| No | Strategy Type | Strategy Name | Reasoning |
|---|---|---|---|
| 1 | Fairness | Shuffling and Mark by Questions | Suggested by a high level experienced participant, personally based on research find it very helpful as well. |
| 2 | Fairness | Double Marking | Suggested by a low level experienced participant, but also found in multiple literature |
| 3 | Fairness + Efficient | Review marks | Suggested by two high level experienced participant, currently already existed a basic system of it, adding some sub-features to make it more efficient |
| 4 | Fairness + Efficiency | Communication | Very common feature, personally find it embedding it in a system would be helpful, saves time for finding contacts |
| 5 | Fairness + Efficiency | Anchor and Compare assignments | Mentioned in both literature and suggested by participants, can also add this feature in the 'non-functional' compare button |
| 6 | Efficiency | Timed Marking | Very basic strategy but does not exist in design |
| 7 | Fairness + Efficiency | Small Criteria- based marking | Suggested by a high level experienced participant, personally thought it would be useful |

Figure F.8: Chris' strategies for fair and efficient marking

**Please note:** Chris' designs were omitted due to space considerations

# Hamdani's strategies for high quality feedback and design screenshots

| Reference | Feature Requirement |
|---|---|
| R1 | Feedback bank shared between markers and the course organiser for that specific course |
| R2 | Collaboration between markers (markers can add a tag/flag asking for another marker's input) |
| R3 | Real Time Collaborative Marking |
| R4 | Student-Marker direct channel of communication |
| R5 | Email reminders for markers to mark well before the deadline for marks |
| R6 | Structuring the feedback into the sections of the feedback sandwich |

Figure F.9: Hamdani's strategies for high quality feedback

**Please note:** Some of Hamdani's designs were omitted due to space considerations

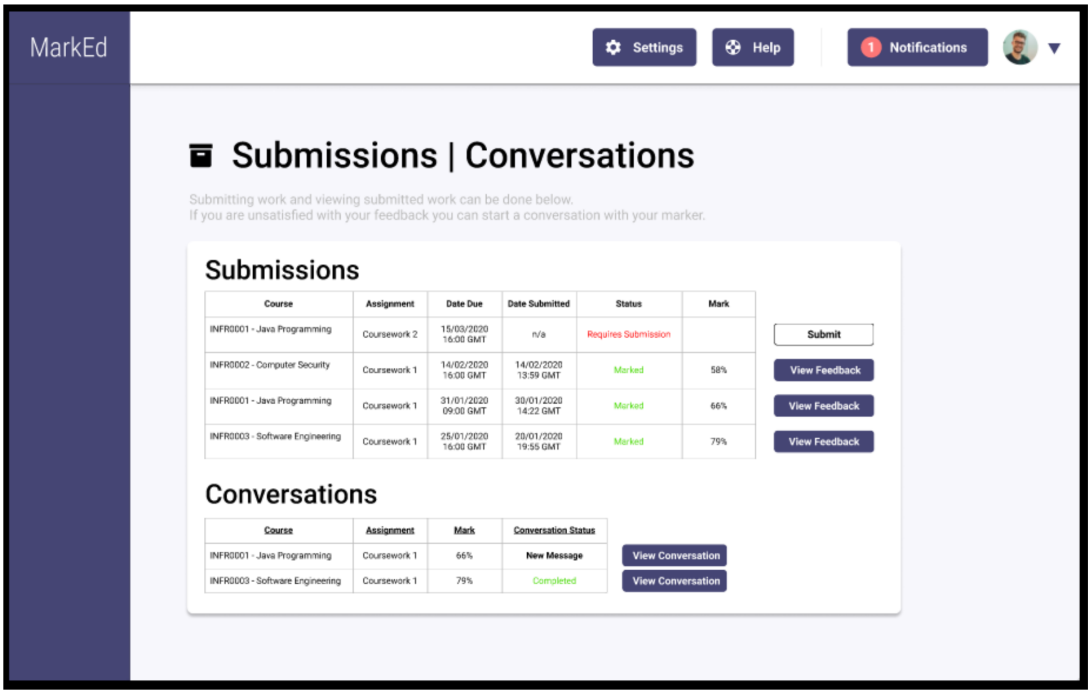



Figure F.10: Student view of the MarkEd Tool: Homepage

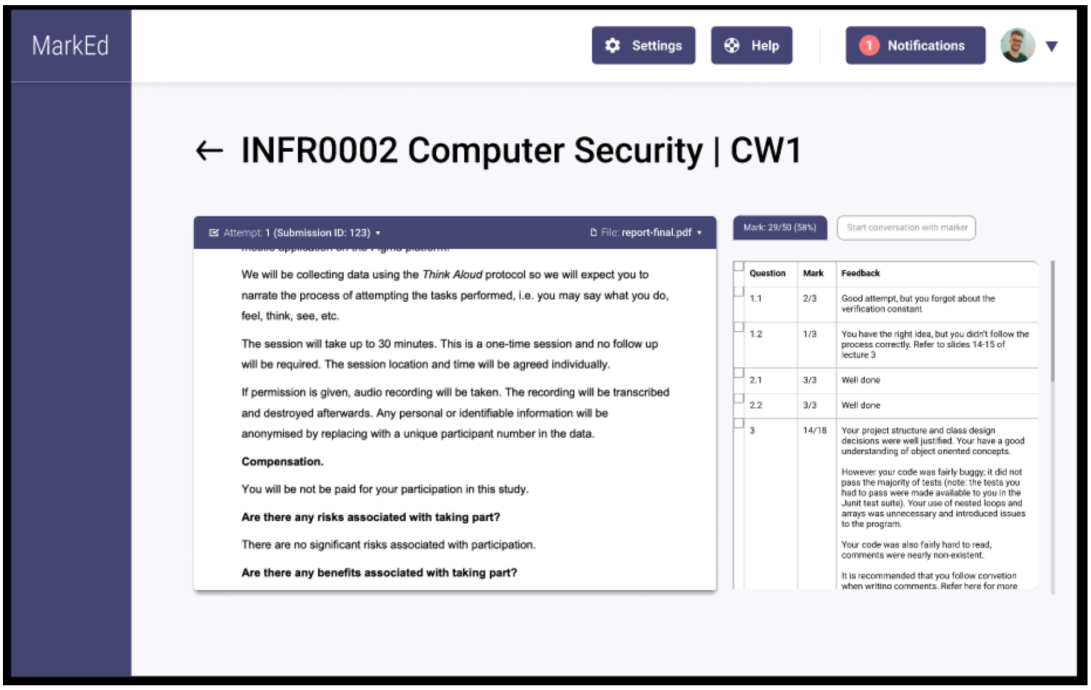



Figure F.11: Student view of the MarkEd Tool: Assignment Feedback

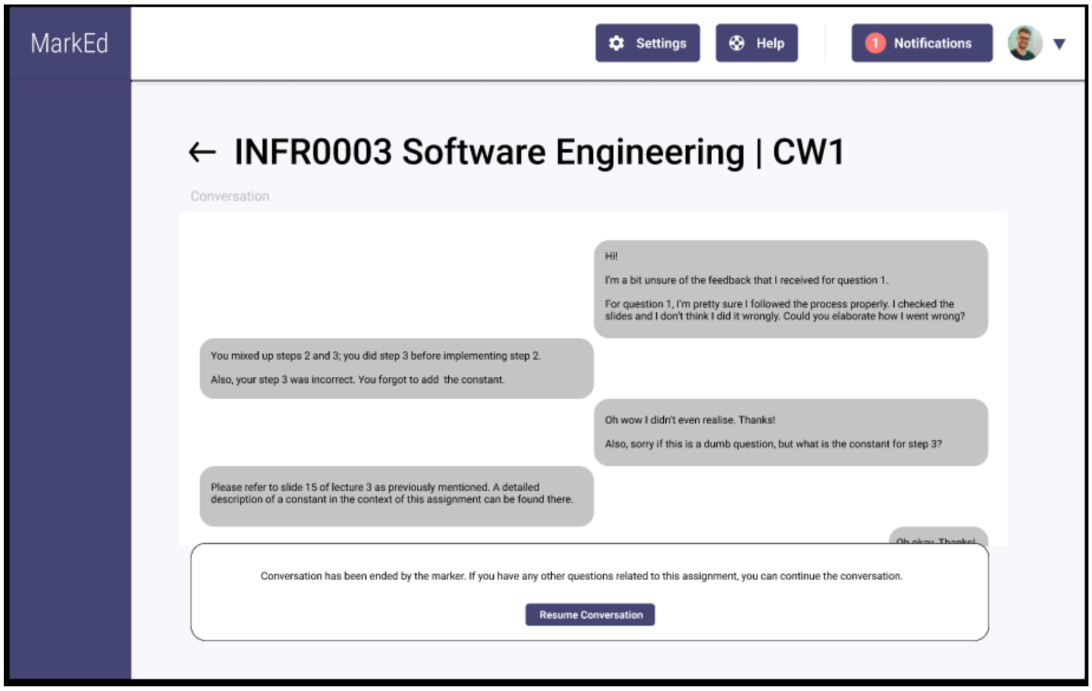



Figure F.12: Student view of the MarkEd Tool: Chat interface (Completed)

# Appendix G

# Xisen's implementation screenshots

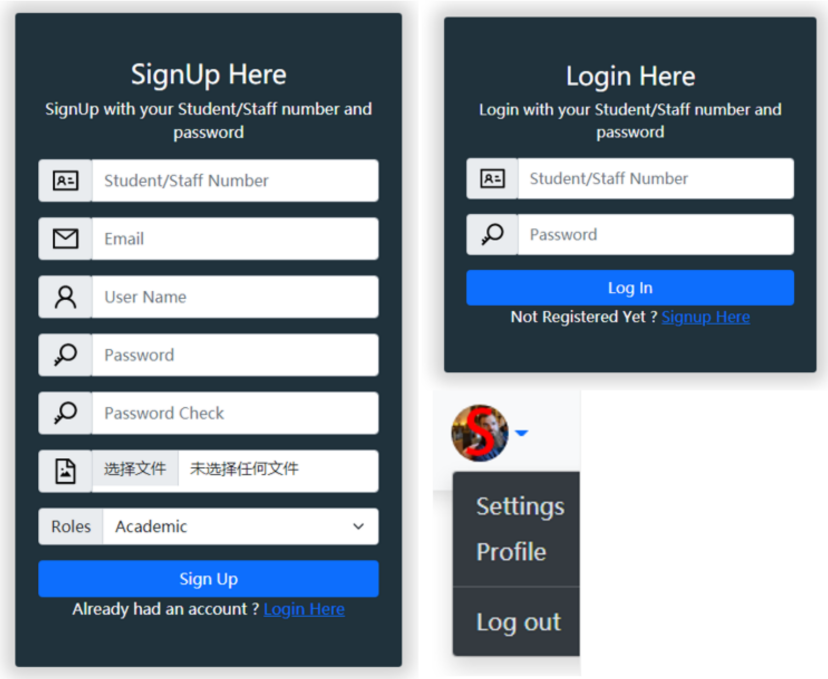



Figure G.1: SignUp, Login and Logout

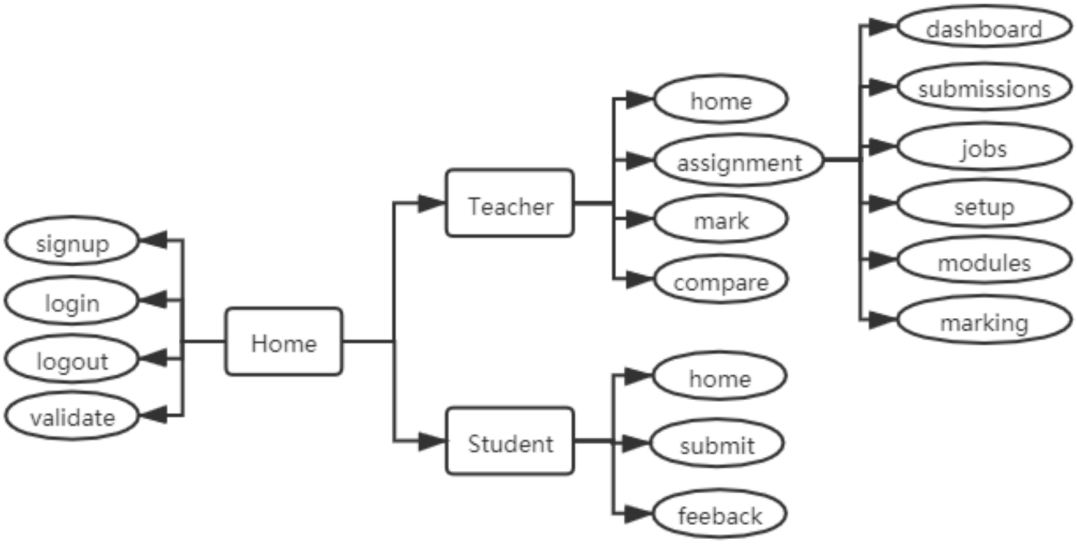



Figure G.2: MarkEd page hierarchy and Router Desgin

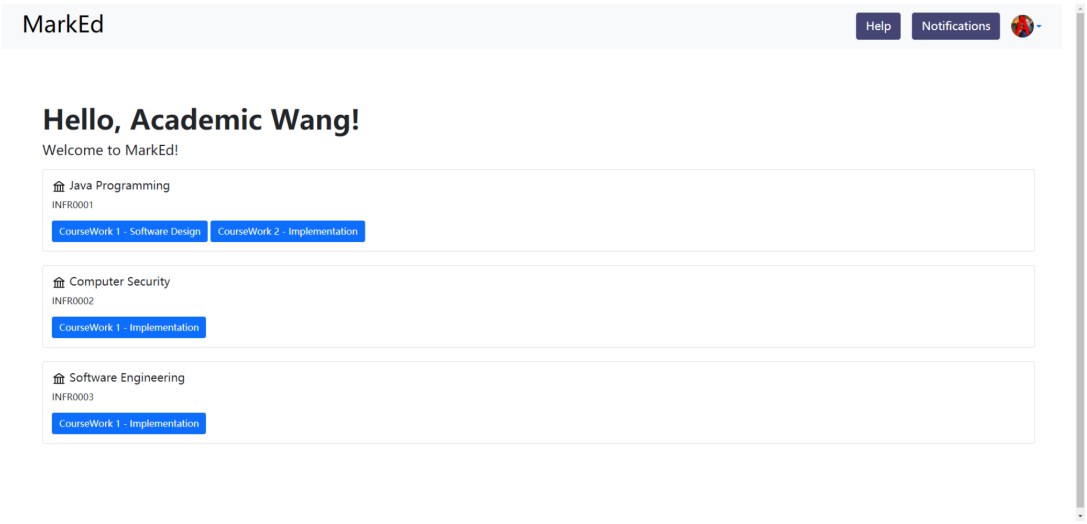



Figure G.3: Teacher's Home Page

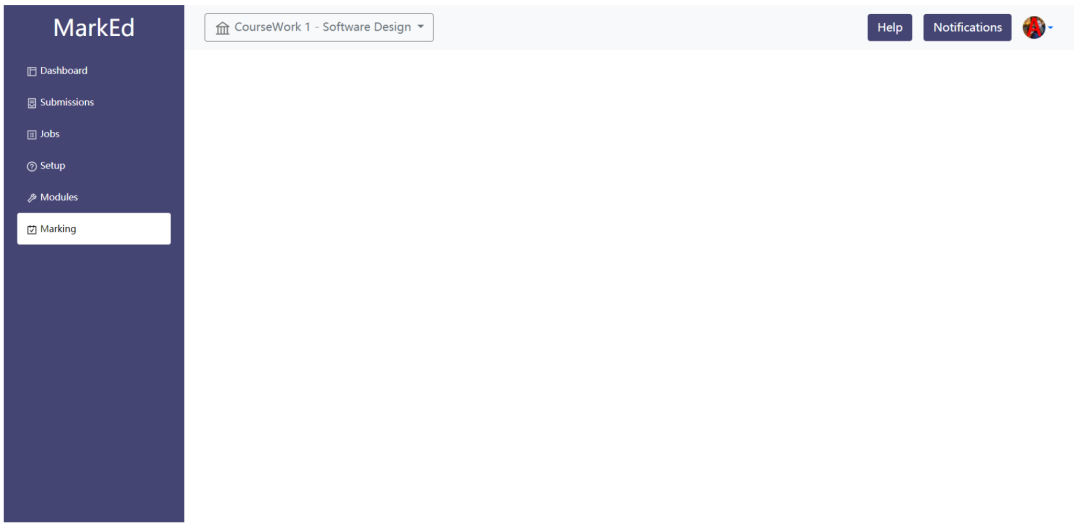



Figure G.4: Template

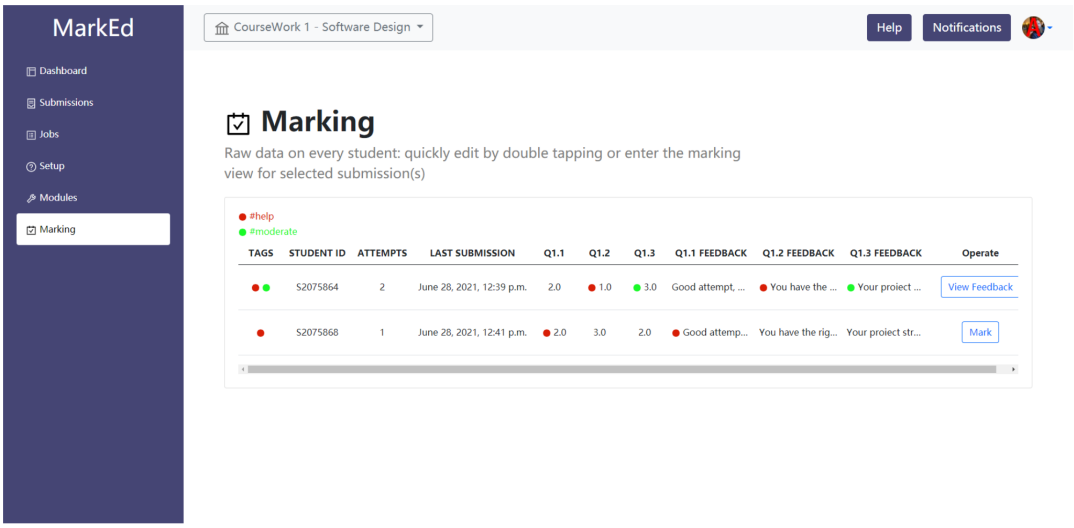



Figure G.5: Marking Data Page

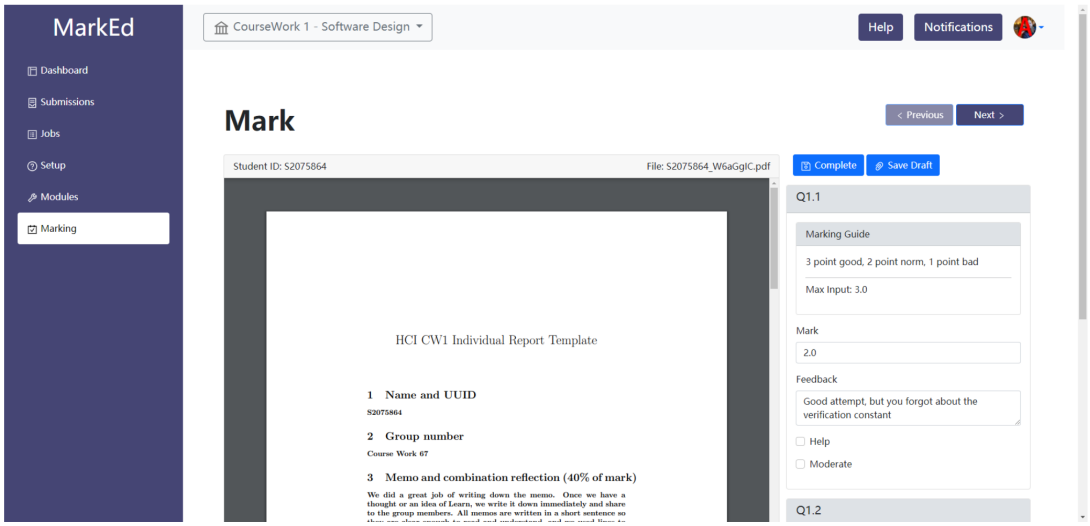



Figure G.6: Mark Page

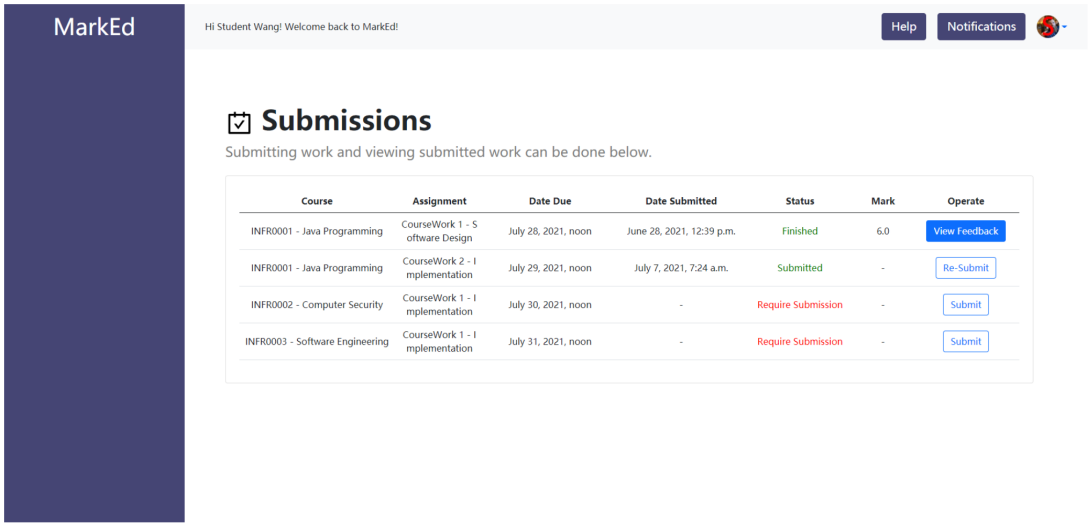



Figure G.7: Student Home Page

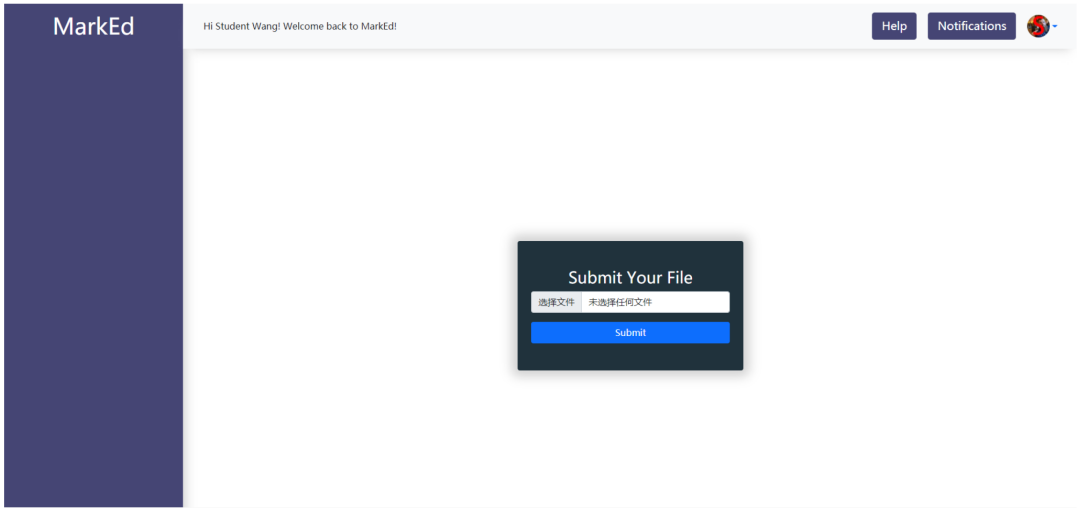



Figure G.8: Submit Page

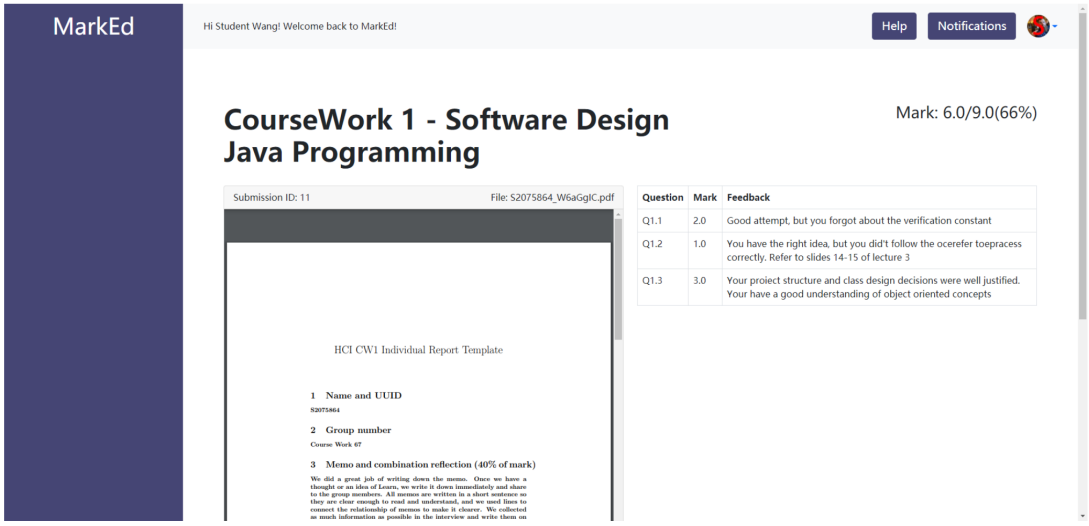



Figure G.9: View Feedback Page

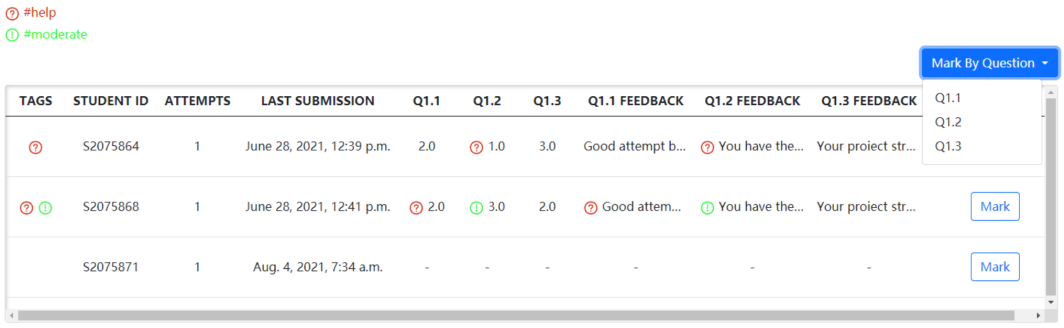



Figure G.10: View All Attempts

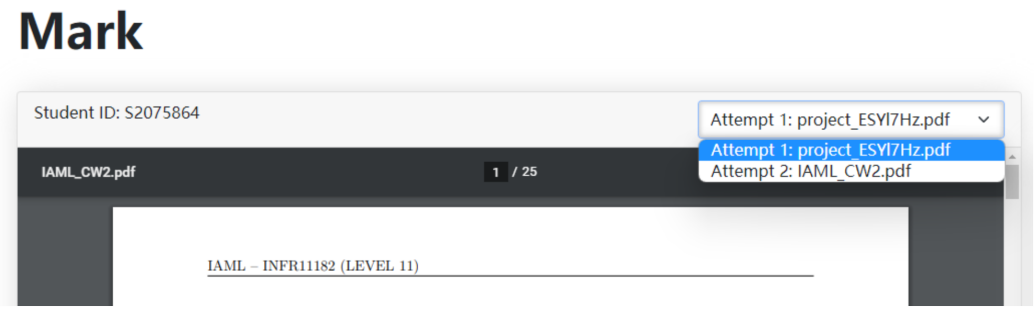



Figure G.11: Mark By Question

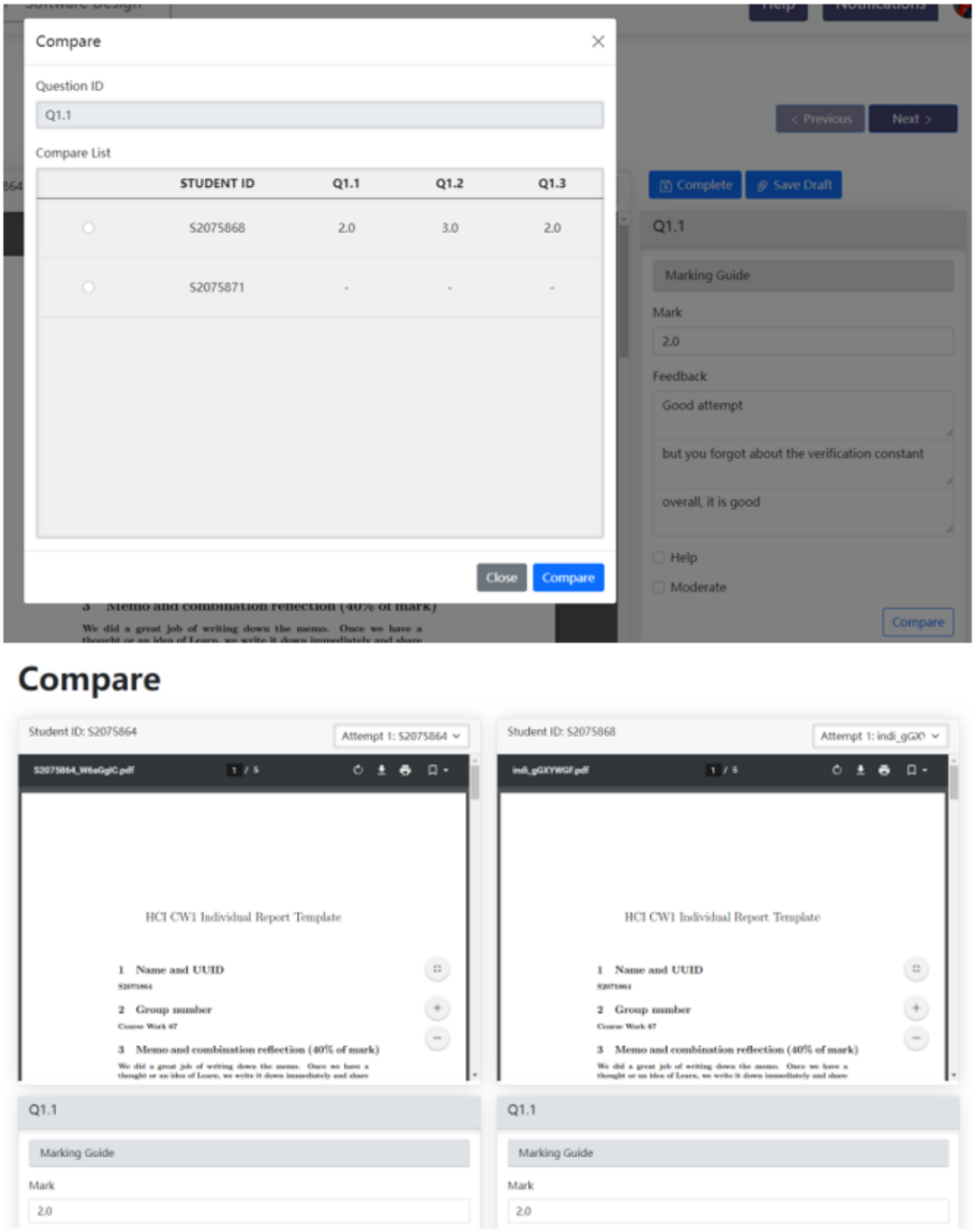



Figure G.12: Compare

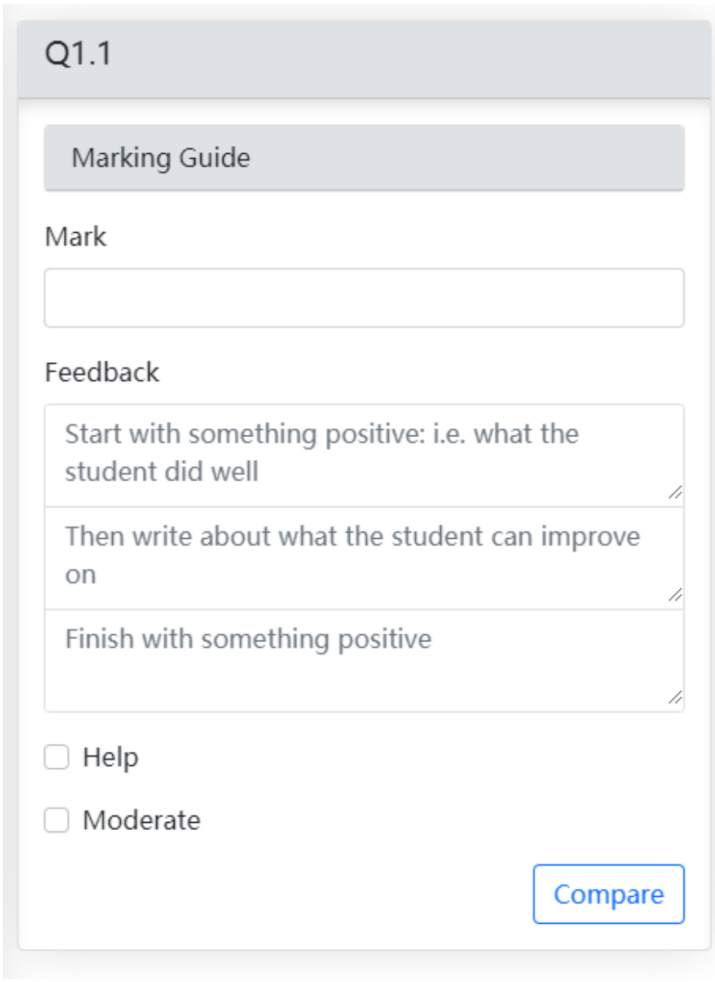



Figure G.13: Feedback Sandwich

# Appendix H

# Xiaofei's implementation screenshots

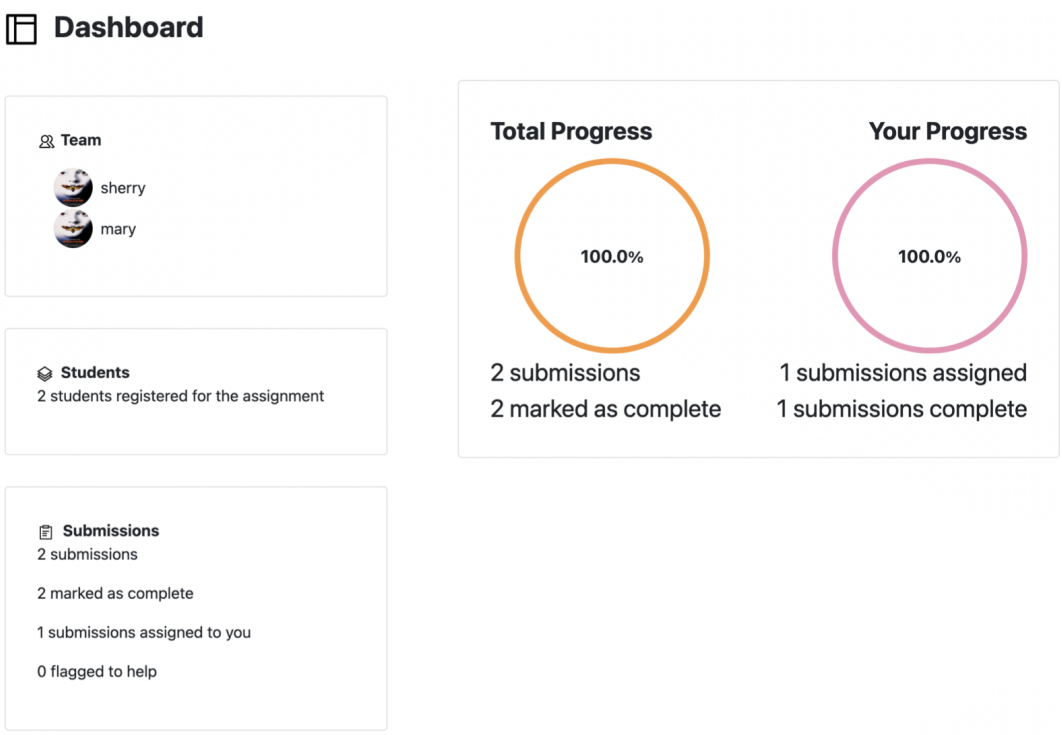



Figure H.1: Dashboard page

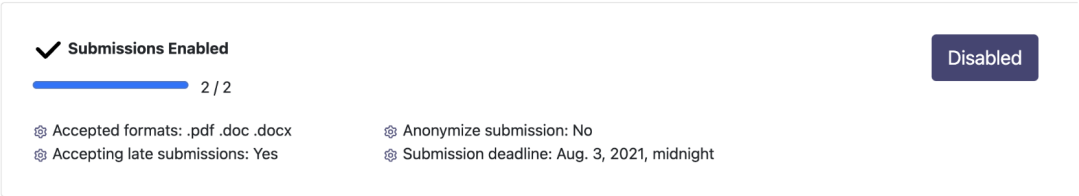



Figure 4.3: The process bar of submissions when student submissions are enabled

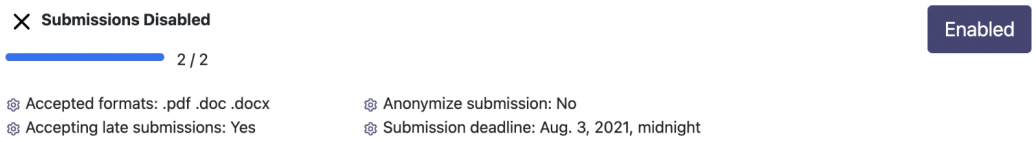



Figure H.2

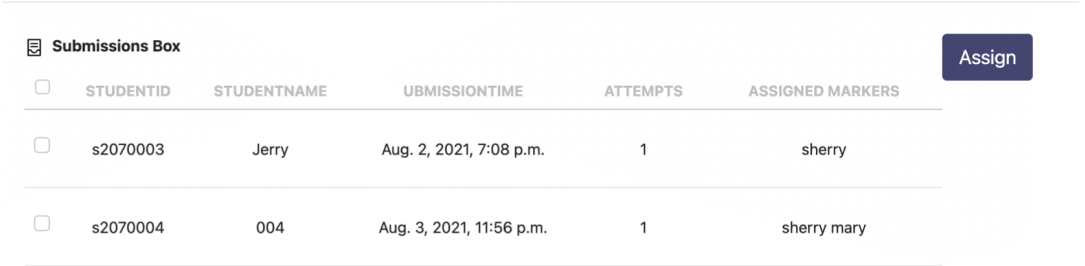



Figure H.3: The table information on the submission page

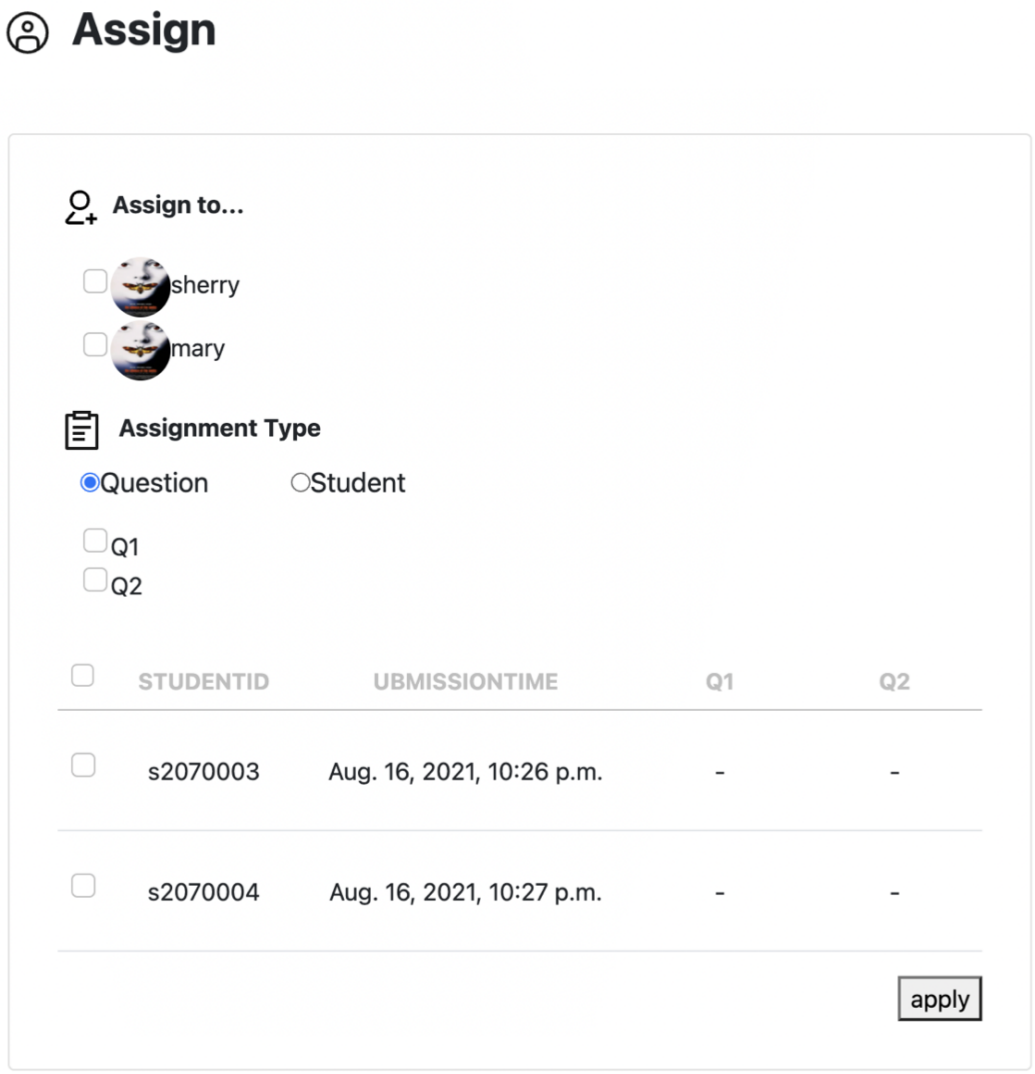



Figure H.4: The assign page

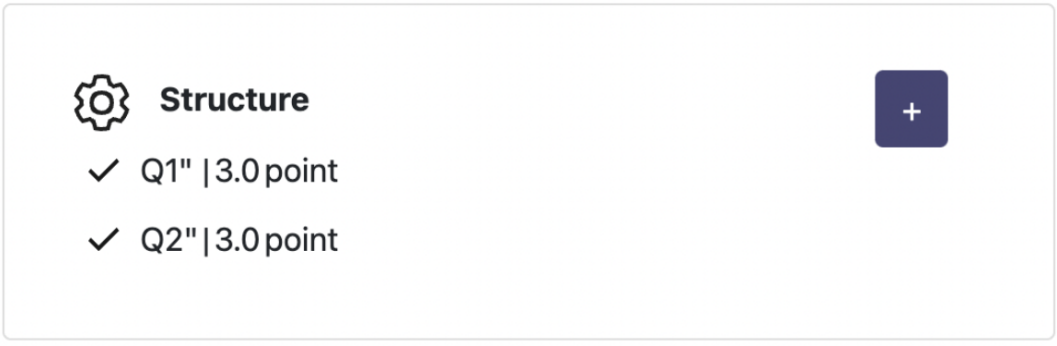



Figure H.5: The structure of assignment

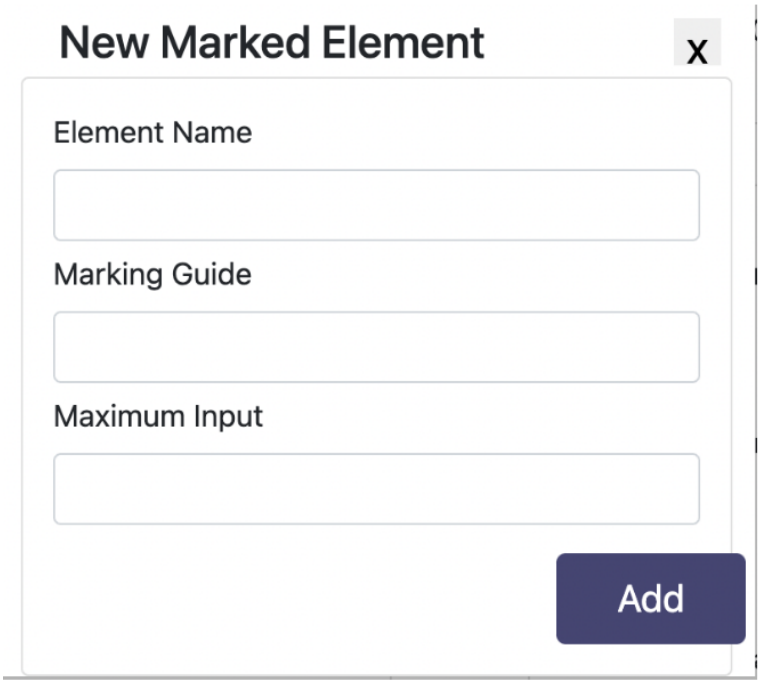



Figure H.6: Add a new element

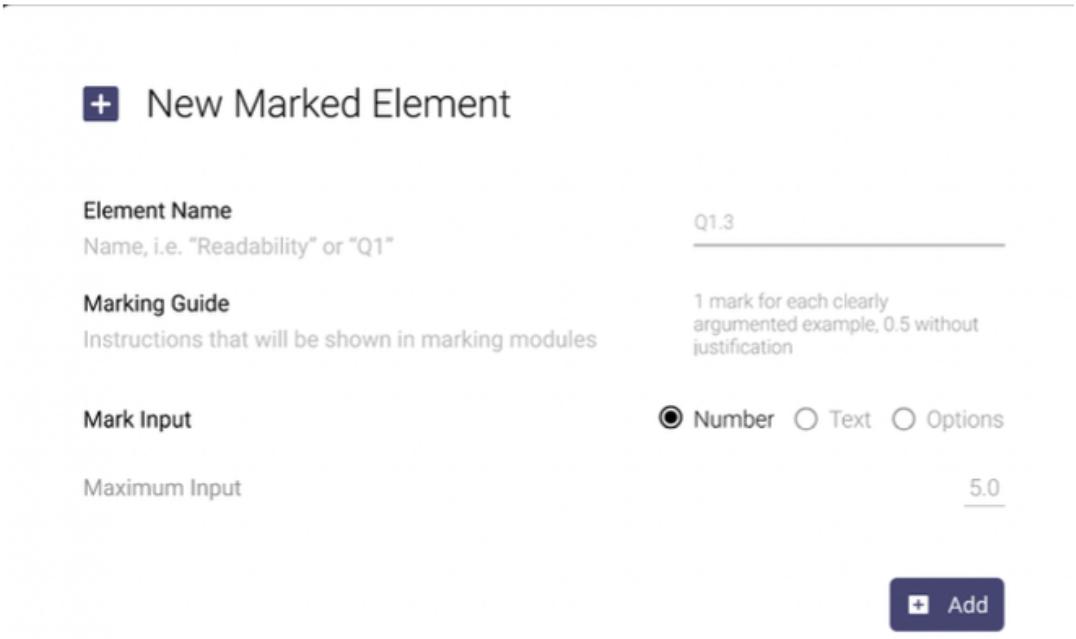



Figure H.7: The original design of the structure part

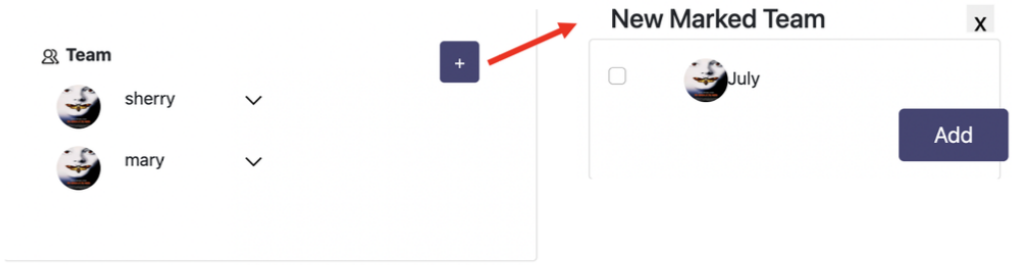



Figure H.8: The process of adding a new member to the team

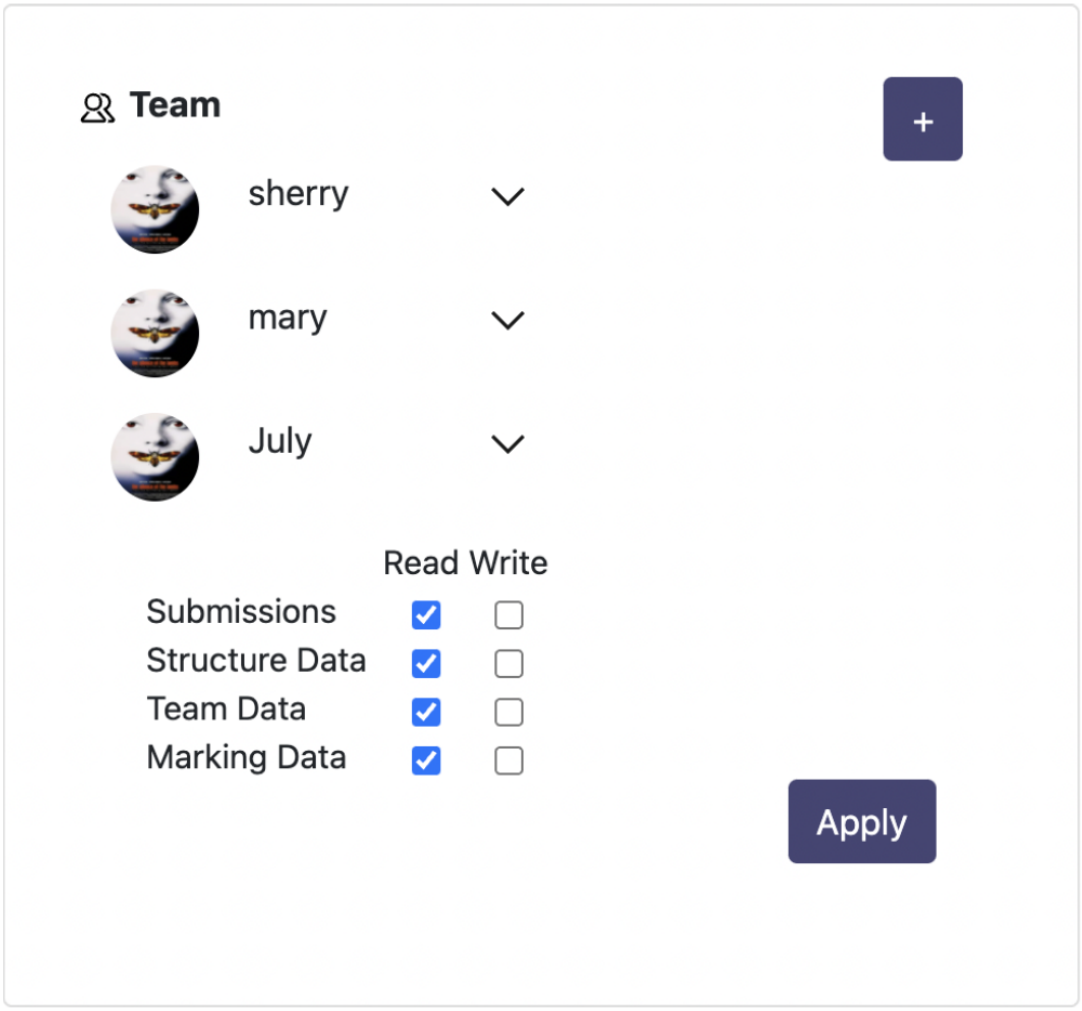



Figure H.9: The process of editing permission of each marker

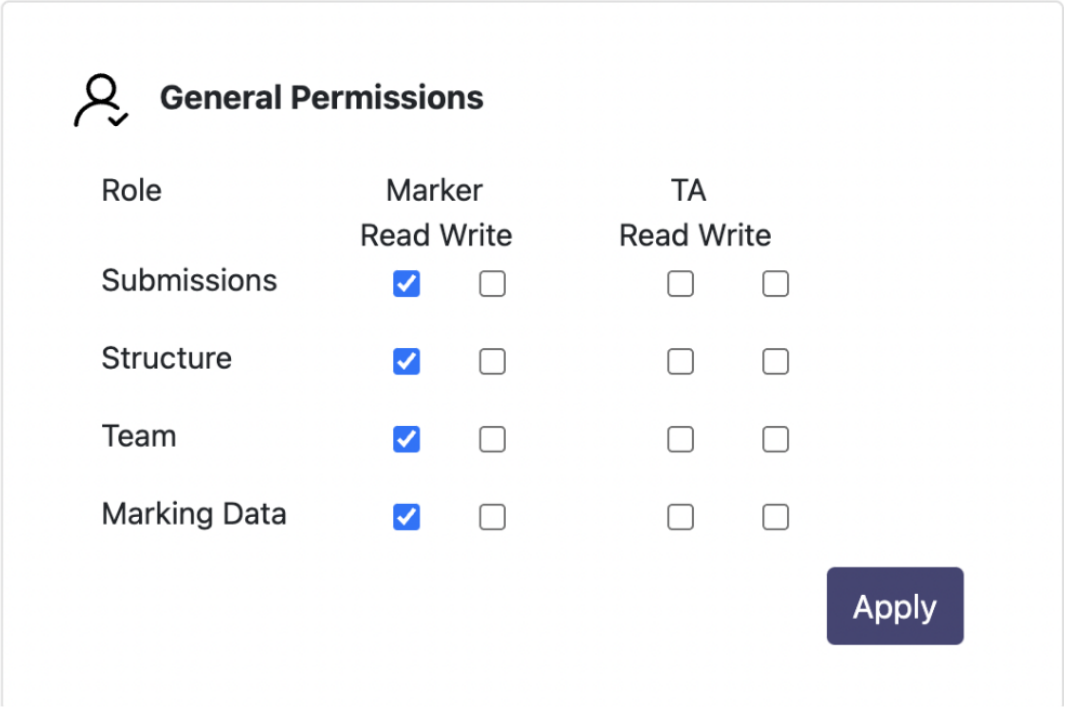



Figure H.10: General permission of all markers

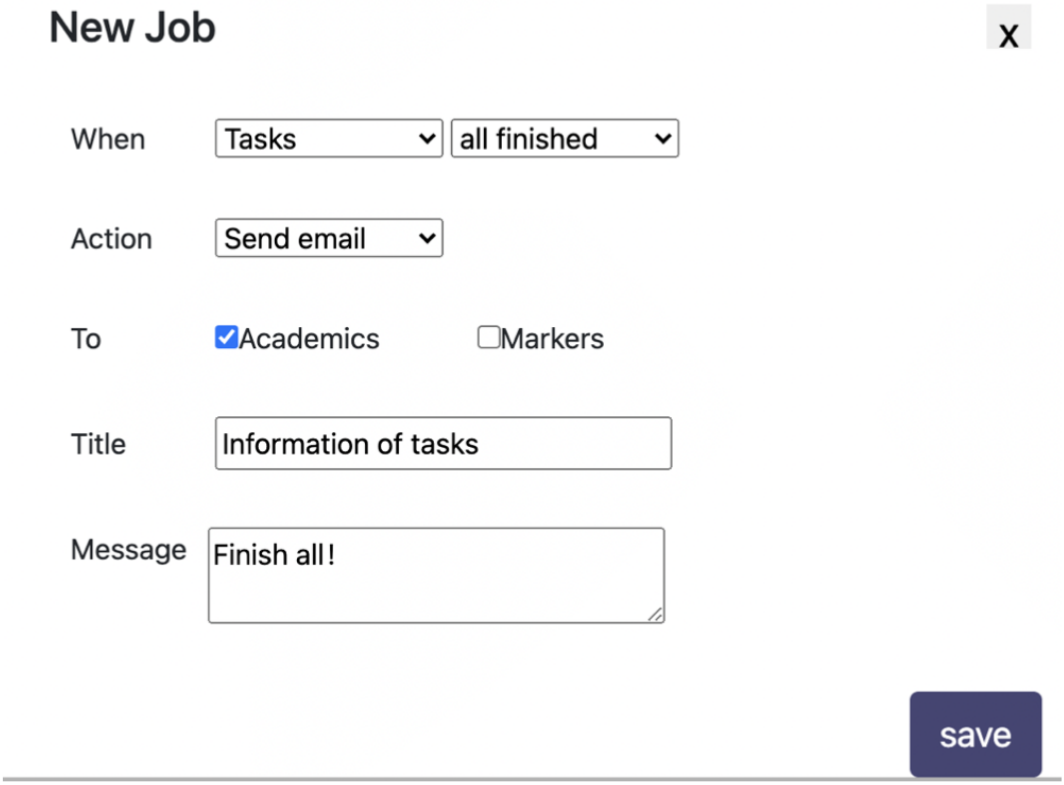



Figure H.11: Add a new job

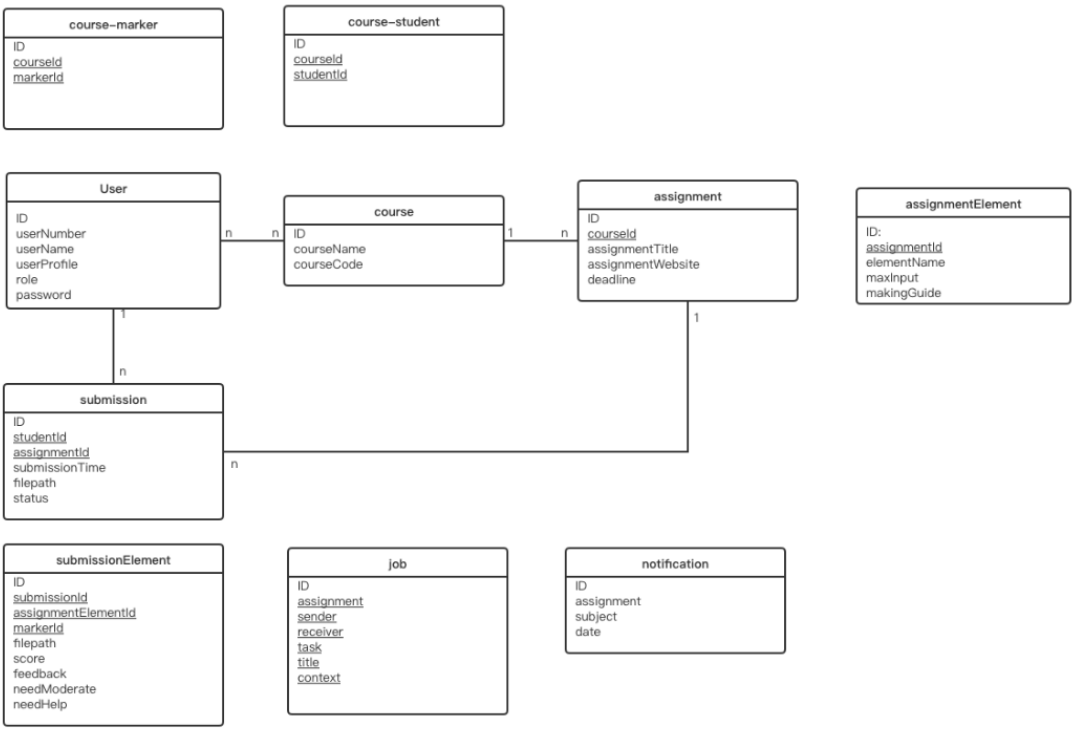



Figure H.12: diagram of database table of the MarkEd