NTNU
Norwegian University of
Science and Technology

Developing a Web Application and a Case-based Reasoning Recommender System to Improve Students' Motivation for Exchange Programs

Lars Liverød Andersen
Truls Mørk Pettersen

Master of Science in Informatics
Submission date: June 2017
Supervisor: Rune Sætre, IDI

Norwegian University of Science and Technology
Department of Computer Science

Abstract

The Norwegian University of Science and Technology has an objective to be internationally outstanding. One of the underlying initiatives of this objective is to increase the number of students that go on a study abroad- or student exchange program. This thesis introduces a prototype created to improve the motivation for students to apply for these programs by recommending courses and universities, and simplifying the course approval process for students. The prototype, named *Utsida*, uses case-based reasoning and the experience reports of previous exchange students to give the recommendations. The Design and Creation research strategy was used to develop and evaluate the two parts of Utsida: a web application part for user interaction and data storage, and a case-based reasoning recommender system part that produces the recommendations. Two methods were used to evaluate the Utsida prototype. The first method was a user study were students at NTNU were able to test Utsida and answer an accompanying questionnaire. The second method tested the case-based reasoning recommender system part in an offline experiment with simulated user queries. The user study showed that Utsida has a very positive effect on students' motivation to apply for a study abroad- or student exchange program. Furthermore, the majority of students also received relevant recommendations for both universities and courses. Furthermore, the significant results of the offline experiment showed a high relevancy on the implemented recommender and supported the results of the user study.

Sammendrag

NTNU har mål om å være internasjonalt fremragende. Et av initiativene til å oppnå målet er å øke antallet studenter som drar på utveksling. Denne avhandlingen introduserer en prototype laget for å forbedre studenters motivasjon for utveksling ved å anbefale fag og universiteter, og forenkle prosessen med å forhåndsgodkjenne fag. Prototypen, kalt *Utsida*, bruker case-based reasoning og erfaringsrapportene til tidligere utveklsingsstudenter for å gi anbefalinger. "Design and Creation" forskningsstrategien ble brukt til å utvikle og evaluere de to delene av Utsida. Den ene delen er en webapplikasjon som håndterer brukerinteraksjon og datalagring, og den andre delen er et case-based reasoning recommender system som produserer anbefalingene. Utsida-prototypen ble evaluert ved å bruke to metoder. Den første metoden var en brukerstudie hvor studenter ved NTNU kunne teste Utsida og svare på et medfølgende spørreskjema. Den andre metoden testet case-based reasoning recommender system-delen i et offline eksperiment med simulerte brukerforespørsler. Brukerstudien viste at Utsida har en svært positiv effekt på studenters motivasjon for å søke utveksling. Flertallet av studentene mottok også relevante anbefalinger for både universiteter og fag. De signifikante resultatene fra offline eksperimentet viste en høy relevanse på den implementerte anbefalingsmetoden og støttet resultatene av brukerstudien.

Preface

This thesis marks the end of two Master's degrees in Informatics at the Norwegian University of Science and Technology (NTNU). The thesis is a joint project between two students in the Artificial Intelligence and Software Engineering fields respectively. The project duration was from August 2016 to June 2017. A main motivation for us to undertake this project was to make the process of applying for an exchange program both easier and more accessible. Therefore, it was very rewarding for us to get positive results from this research project. We hope our project will contribute to more exchange students from NTNU in the future, and that it may increase the interest for further work on the subject.

We would like to thank our supervisor, Associate Prof. Rune Sætre, for the useful guidance, encouragement, and feedback he has provided us with throughout the year and Assisting Prof. Terje Rydland, for co-supervising the project and participating in interviews. We would also like to thank all the staff at the NTNU Department of Computer Science who assisted us in the process. In particular, we would like to thank Associate Prof. Kerstin Bach for assisting us with everything regarding CBR and myCBR and to Prof. Agnar Aamodt for helping us validate our initial project idea. Furthermore, we would like to thank the students who participated in usability tests during the development process and those that used their time to answer our questionnaires. Finally, we would like to thank our family and friends for the continuous support and encouragement.

Lars Liverød Andersen & Truls Mørk Pettersen

Trondheim, 08.06.2017

Table of Contents

List of Tables

List of Figures

Abbreviations

AI	=	Artifical Intelligence
API	=	Application Programming Interface
CBR	=	Case-Based Reasoning
CBRS	=	Case-Based Reasoning System
CBR-RS	=	Case-Based Reasoning Recommender System
CSV	=	Comma-Separated Values
HTML	=	HyperText Markup Language
IDI	=	Department of Computer Science at NTNU
IS	=	Information System
JSON	=	JavaScript Object Notation
NTNU	=	Norwegian University of Science and Technology
OIR	=	The Office of International Relations at NTNU
REST	=	Representational State Transfer
RQ	=	Research Question
SDK	=	Software Development Kit

Chapter 1

Introduction

This research project was initiated to make it easier for students at NTNU to apply for a study abroad- or student exchange program by using a prototype, named Utsida, as a proof of concept to replace parts of the current manual process. This chapter elaborates the motivation and background for that initiation. It also details the research questions and goals for the solution, gives a brief overview of the research methodology, and describes the structure of the thesis.

1.1 Motivation and Background

Studying abroad can be an important way to improve intercultural communication skills that are increasingly more important in international businesses and for cooperation between different cultures [38]. A multicultural experience is also shown to enhance creativity [19] and improve the career opportunities for students [10]. Because of the beneficial effects of multicultural experiences, The Norwegian University of Science and Technology (NTNU) has set the following goal of international mobility for degree students: “By 2017, at least 40% of NTNU’s degree students should have a period of study at an edu-

cational institution in a foreign country lasting one or two semesters"[1]. Norway has also committed to the Bologna Process[2], with one of the goals being that 20% of the students completing a study in Europe should undertake a minimum of three months of their education in a different country by 2020.

With only 25% student mobility at NTNU in 2015 [25], NTNU has to, to reach their goal of 40% international mobility, look at ways to increase motivation for students to do a study abroad- or student exchange program. One of the ways could be to digitize the current process of applying for a study abroad- or student exchange program; this could increase student motivation and reduce the workload for advisers. The difference between an exchange program and study abroad program is that a student exchange program involves students exchanging places and the student only pays tuition fees at their home university. In a study abroad program, however, the student is enrolled in the abroad university and pay the tuition fee. Both of these programs have similar application procedures at NTNU, and will, henceforth, be referred to as an exchange program. Before applying for an exchange program at NTNU, students have to find and approve courses to take at the university abroad. This process of approving and finding courses is mostly done manually by both advisers and students. It could be replaced by an information system (IS) that simplifies the process and consequently increase students' motivations for exchange.

1.2 Personal Motivation

Replacing tedious manual work with an intuitive and easy to use IS is something that motivates both of the authors. The project was first proposed by our supervisor, Associate Prof. Rune Sætre, as a study exploring the similarity between courses from different universities. However, due to several ongoing similar studies, we instead proposed to focus on finding ways to improve the exchange course approval process. One of the authors, Truls Mørk Pettersen, has previously done an exchange program where he experienced that the current process has room for improvement. He therefore wanted to study ways to increase

[1] NTNU's international action plan: `http://www.ntnu.edu/international-action-plan`
[2] Bologna Process: `https://www.ehea.info/`

the motivation and reduce the workload for students.

1.3 Research Questions and Goals

The primary goal of this project is to create a prototype, named Utsida, that increase the motivation for students at NTNU to apply for an exchange program. To achieve this goal the prototype should simplify the exchange course approval process and give relevant recommendations on courses and universities. The second goal is to use the popular case-based reasoning (CBR) methodology for producing the recommendations. The following list summarizes the research goals.

- Goal 1: Create a prototype that improves the motivation for students at NTNU to apply for an exchange program.
- Goal 2: Use the CBR methodology to give relevant recommendations on exchange program universities and courses.

Based on these goals, and the knowledge gained in the preliminary research (Ch. 2) the specific research questions this study seeks to answer are:

- **RQ1:** What effect would an information system for recommending and assisting exchange course selection have on students' motivation for doing a study abroad- or student exchange program at NTNU?
- **RQ2:** How suitable is the case-based reasoning methodology for recommending relevant universities and courses for a study abroad- or student exchange program?

1.4 Methodology

Oates' [26] model of the research process, detailed in Chapter 4, was used to define the chosen research methodology. The research questions were developed from motivation (Ch. 1), experiences (Ch. 2) and related work (sec. 2.3). The Design and Creation research strategy was used to create Utsida, a prototype system with two parts: a web

application and a cased-based reasoning recommender system (CBR-RS). These systems were, in turn, evaluated through questionnaires which produced the data needed to answer RQ1 and partly RQ2. To fully answer RQ2, and further quantify the result of the questionnaires, an offline experiment was performed on the CBR-RS to analyze whether it yielded relevant recommendations. See Chapter 4 for a full description of the research methods used.

1.5 Thesis Structure

The thesis is structured in eight chapters. Firstly, Chapter 1 and 2 answers why this research was conducted by exploring the motivation, problem context, and related work. Chapter 3 then introduces theory on essential concepts identified in the preliminary research. Furthermore, Chapter 4 and 5 answers how the research project was conducted by detailing the chosen research methods and the implementation of the prototype. Chapter 6 shows the results produced by the research methods, while chapter 7 interprets these results in regards to the research questions. Finally, Chapter 8 ends the thesis by concluding the work done and presenting ideas for future work.

The following paragraphs further detail the content and purpose of each chapter.

Chapter 2: Preliminary Research Describes the preliminary research that led to the production of the research questions and the goals of this project. This includes knowledge attained on the problem domain, review of the current approach, and a summary of related work found through literature review.

Chapter 3: Theory Presents the theory and information about concepts and tools which were essential for creating the prototype needed to answer the research questions. The main focus lies on CBR and recommender systems.

Chapter 4: Research Method Describes the methodological choices of this thesis. This includes presenting the overarching research strategy: Design and Creation, used as a guideline to develop the prototype. Furthermore, the chapter details the data collection methods and data analysis used to evaluate the prototype and consequently answer the research questions. These methods are questionnaires, offline experiments, and quantitative data analysis. Finally, important ethical and practical issues that influence the project are elaborated.

Chapter 5: Implementation Elaborates on the architecture and implementation of the prototype, Utsida. Including decisions made for design, frameworks, and technologies, as well as how the CBR-RS was modeled and implemented.

Chapter 6: Results Presents the key results acquired with the utilized research methods on the prototype. These results are represented by diagrams and tables.

Chapter 7: Discussion Includes a comprehensive discussion were the results are interpreted in regards to the research questions, and possible limitations of the research project are presented.

Chapter 8: Conclusion Concludes the thesis based on the interpreted results and presents ideas for future work on the subject. These ideas includes new features for the prototype, extensions to CBR-RS, and ways to evaluate the system in real-life conditions.

Chapter 2

Preliminary Research

Before producing the research questions and deciding on the specific research methods to be used, it was important to understand the problem domain and related work done in the field. This chapter details that process and the knowledge gained through interviews, meetings and literature review. The chapter is concluded with the identified system requirements and research gap.

2.1 Interviews

Interviews and informal meetings were conducted to understand the problem domain and gain information on the current approach of applying for an exchange program from NTNU. Two advisers from the Computer Science Department at NTNU (IDI) were interviewed, and a meeting was conducted with the Office of International Relations at NTNU (OIR). These engagements were also conducted to gain valuable feedback on possible difficulties and limitations that might be encountered in the project.

2.1.1 Student Advisers

Student advisers were interviewed to gain a better understanding of their role in the process of approving courses for an exchange program. In the context of this process, student advisers mostly work on reviewing courses found by students and gives recommendations on courses to take, based on the students' study plan. Advisers also give information to students on partner universities and on the process as a whole. Several problem areas were noted by the advisers. Firstly, a high workload in peak application periods makes it difficult to spend enough time reviewing students' courses. Secondly, a generally large number of students makes it difficult for a personal follow-up throughout the exchange program duration. Finally, the current process of approving exchange courses is arduous. The advisers were highly positive to a research project which could find ways to resolve some of these issues. Features suggested by advisers included digital course applications, digital lists of approved courses for each department, and automatic expiration of course approvals in this list based on approval date.

2.1.2 The Office of International Relations

The main reason for meeting with the staff of the OIR was to gain insight into the current process of applying for exchange programs at NTNU. Another reason was to gain access to the exchange experience report database that OIR maintains. These reports would be highly valuable for giving the recommendations in the prototype. The database access was, however, found to be restricted because it stores sensitive personal information. OIR instead gave permission for this project to extract the reports directly from the website. The OIR also presented, during the meeting, the results of a study on student motivation they conducted in the spring of 2016 that were useful for determining focus areas of the project. Furthermore, they noted some limitations the project might encounter, such as different practices at various departments and little information on unpopular exchange locations. The OIR suggested many features, among these were support for differences in study points of courses, and recommendation of courses and universities at locations with fewer students.

2.1.3 Previous Exchange Students

Because one of the authors has previously been on exchange, some of the difficulties of the current process were already identified. It was, however, useful to interview students from different departments and with other experiences to gain a better understanding of the problem domain. The focus of the interviews was to ask questions on issues with the current process and to brainstorm possible improvements. Significant issues noted by the students were: concern of courses not being approved after an exchange period, difficulty changing courses during the exchange period, and the tedious manual process before an exchange period. The proposed features from students were: A map of where other students have done an exchange program, combining courses from NTNU with possible replacements at a foreign university and recommendation of universities and courses based on the preferences of each student.

2.2 Current Approach

The following description is based on the current approach at the Computer Science Department (IDI) at NTNU. It may, therefore, be some differences at other departments. See Figure 2.1 for a short overview of the whole application process. The boxes colored in green are the targeted problem areas of this research project.

Students are first required to find and choose the university where they want to do an exchange program. There is a considerable number of possible choices, and NTNU only gives information on universities with an cooperation agreement. After the university is chosen, the students have to find courses that can be approved as a replacement for the courses originally planned at NTNU. Most of the departments do not have a system where students can enter their replacements or find the previously approved ones. One exception is IDI, which utilize a digital table to make it easier for students to find course replacements. However, manually maintaining the table is tedious work and it lacks useful features such as data validation.

[1] NTNU's application process: `https://innsida.ntnu.no/utenlandsstudier`

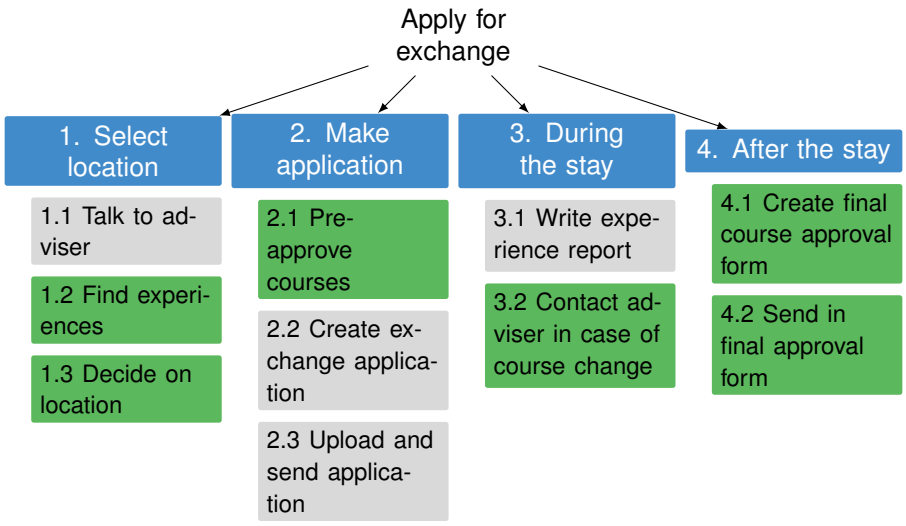


Figure 2.1: Current approach to apply for an exchange program at NTNU[1]

Throughout the duration of an exchange program, the students often have to change their courses due to unforeseen circumstances, making the original approval invalid. To gain a new approval, a student usually has to send an email to the adviser of their department to get the new courses approved. If this process takes a while, and the course registration due date is close, the student may have to take a course without approval from NTNU, risking not getting it approved at the end of the exchange program.

2.3 Related Work

Mazzarol and Soutar [22] found that students' general motivation is influenced by the amount of available information on a university and its courses. Among the several factors which were reviewed, the "knowledge and awareness" factor proved to be the most influencing one for choosing an international study location. Therefore, when researching how an IS can improve the motivation of students to participate in an exchange program, focusing on providing knowledge and awareness is important.

Furthermore, the OIR conducted a study on 464 exchange students in 2016 to investigate

what factors could help increase the motivation to apply for an exchange program [27]. Among the 464 participating students, 84% answered that a list of previously approved exchange courses could increase the number of students that go on exchange, and 80% answered that pre-approved course packages could increase the numbers. These results indicates that an IS which target these areas could be beneficial to increase the number of students that apply for an exchange program.

Several approaches have been made to replace the manual process of choosing courses with an IS. One example is by using a decision support system that advises students on their course selection based on their program requirements and the course's prerequisites, as done at the University of Dhaka [31]. Furthermore, student course recommendation can also be done intelligently by using recommendation engines and data mining techniques to give relevant results. Sherpa [9] is a system that has been made especially for the goal of providing course recommendations. Data mining techniques can also be used to predict whether a student will fail or pass on a course [37].

Recommender systems are used extensively in both research and commercial products. These systems provide several possible solutions that might be viable for the user to choose. Examples of recent published papers on different recommender systems are: T-Finder [41]; a system which recommends locations for taxis and passengers for the most efficient trips, iTravel [40]; a peer-to-peer recommender system for exchanging ratings on different tourist attractions, and Happy Movie [29]; a Facebook application for recommending movies to groups. When designing a recommender system it is important to know which data that should be used as attributes. Cubillo et al. [21] identified the different decision making factors for international students and their grouping. The five main groups were city effect, country image effect, personal reasons, institution image and program evaluation. One way to implement a recommender system is by utilizing case-based reasoning (CBR).

CBR is today both a recognized and well-established method in several fields of sciences, such as health science [4]. Because CBR is proven to be a suitable methodology for complicated problems where uncertainty is involved [30], it has been applied to problems

such as diagnosing chronic diseases in combination with data mining [17], and finding the most suitable study program at a senior high school [23].

Case-based reasoning recommender Systems (CBR-RS) are also commonly used in decision problems, typically when there is not one exact answer, but rather several different good answers. Many types of problems can be solved using CBR-RS systems. Examples of such systems are: MOOC-Rec [8] a system which finds the most suitable Massive Open Online Courses for e-learning, DieToRecs [14] which recommends travel destinations and activities, and helps planning trips, and Entree [34], which recommends restaurants based on a users preferences such as price, type of food and atmosphere.

2.4 Conclusion

This section presents the system requirements for the prototype and the identified research gap the project targets.

2.4.1 System Requirements

The different needs and possible contributions identified in the preliminary research through literature review, interviews and analyzing the current approach are shown in the requirements list (Table 2.1). In the survey [27] sent out by the OIR to students that have been on exchange programs, 85% of 464 students stated that a list containing earlier approved course matches would increase the number of students going on study exchange programs. This functionality, therefore, felt especially important to include as a requirement. The findings were also used to help define the research goals and questions (sec. 1.3).

2.4.2 The Identified Research Gap

Among the many applications utilizing the CBR methodology, none has covered the issue of finding the most suitable country, university and courses when going on an exchange

program. This gap in knowledge combined with the needs identified in the system requirements is what this project will target through the research questions and goals (sec. 1.3).

Table 2.1: System requirements derived from preliminary research

#	Requirement	Priority	Source
1	Recommend universities and courses	1	Students, OIR
2	List previously approved course matches	2	Students, advisers
3	Support matching of exchange courses with NTNU courses	2	Students
4	Handle course approval applications	3	Advisers
5	Show number of exchange students that have been in each country with a world map	4	Students

Chapter 3

Theory

This chapter presents the theory on CBR methodology and CBR Recommender Systems (CBR-RS), which were used to create Utsida. It also introduces myCBR; the software tool used to model and implement the CBR-RS part of Utsida.

3.1 Case-based Reasoning

Case-Based Reasoning (CBR) is a methodology in the field of AI for capturing and reusing experiences. The CBR methodology use these experiences to yield solutions to a large range of problems where uncertainty is involved [30]. Implied by the name, CBR can be broken down into two main concepts: *cases* and *reasoning*. A case is an experience of a solved problem and consists of two parts: a problem description and its solution. Reasoning refers to the approach of finding solutions to problems by evaluating different experiences (i.e. cases). Reasoning in CBR differs from other kinds of reasoning because it does not lead from true assumptions to true conclusions [30]. This means that for two cases with similar problem descriptions, the solution to one might not be the solution to the other. The recorded experience in the first instance may not be exactly similar to the other case.

3.1.1 Case Representation

A case is usually represented as a feature vector, consisting of pairs of attributes and their values. An example is a diagnosis of a sick patient, as shown in Table 3.1. Here, the attributes of the case are a collection of symptoms, and the values are whether the patient has the symptom or not. The solution to the case is the final diagnosis of the patient. A *Concept* represents a case and has a case-representation, global and local similarity measures, and a weight value for each attribute in the case-representation.

Table 3.1: Case representation example

Part	**Attribute**	**Value**
Problem	Nausea	Yes
	Fever	Yes
	Malaise	Dizzy
	Blood pressure	Normal
	Vision changes	No
	Shortness of breath	No
Solution	Diagnosis	Influenza

3.1.2 The CBR-Cycle

Aamodt and Plaza [1] defined a model which explain the problem solving cycle in CBR (Fig. 3.1). The cycle defines the four distinct steps, the 4 R's, of the CBR process and can be summarized as followed.

1. Retrieve: The most similar case is retrieved from the system's case-base so that it can be compared to a new case.

2. Reuse: Solutions are adapted to fit the new case.

3. Revise: The new case is evaluated and tested for success, and is repaired if it fails.

4. Retain: The new case with the given solution is retained in the case base for future usage.

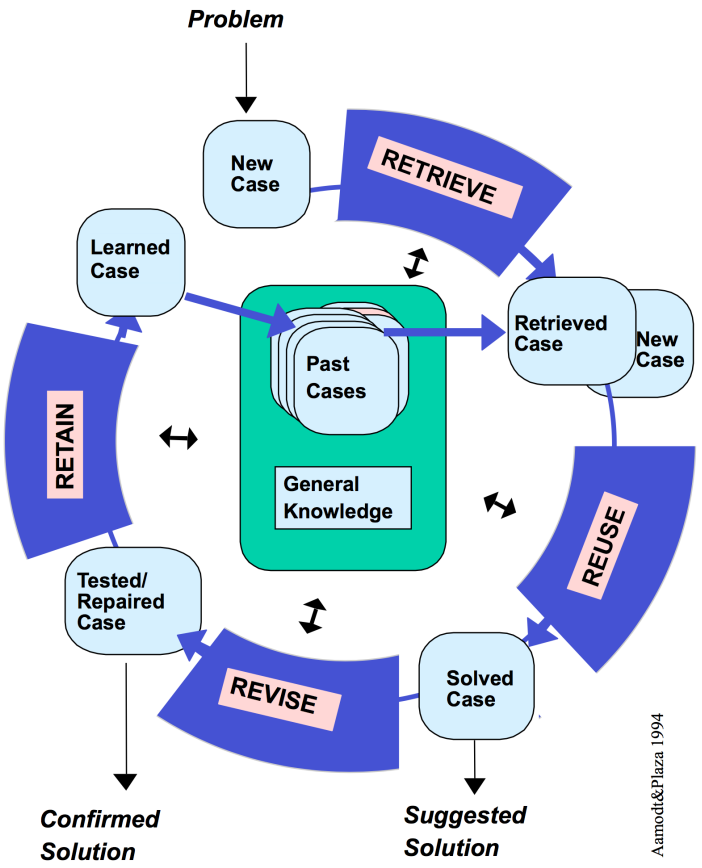


Figure 3.1: The CBR Cycle (Aamodt & Plaza) [1]

3.1.3 Case-based Reasoning Systems

A case-based reasoning System (CBRS) is a system which uses the CBR methodology to solve problems. When the system receives a new problem, it first performs step 1 in the CBR-cycle (retrieve) by selecting the most similar case. This step includes using similarity measures to determine the similarity between each case and the new problem. Next, the system adapts the solution of the case to the new problem (step 2, reuse). The proposed solution to the new problem is then tested by the system, and if found necessary, repaired (step 3, revise). Finally, the adapted solution is stored as a new case in the system's collection of cases, called the case-base (step 4, retain).

3.1.4 Similarity Measures

Similarity measures calculate the degree of similarity (i.e. similarity score) between an attribute in a problem and the corresponding attribute in a case from the case-base. The local similarity measure for each attribute is usually different and is often tailored to suit the data type of the attribute, e.g. integer, string or symbol. Each attribute in a case also has a weight value, which decides the importance of that attribute. The global similarity between a case and a problem is calculated by using an amalgamation function such as weighted sum. A weighted sum function use the local similarity scores with the weights as coefficients to calculate the global similarity.

3.2 myCBR

myCBR[1] is a tool for rapid prototyping of CBRSs and focuses on the retrieval step of the CBR cycle [33]. It is developed and maintained as a joint project between the German Research Center for Artificial Intelligence and the School of Technology and Computing at the University of West London. The myCBR project includes two parts: a Software Development Kit (SDK), and a graphical user interface (GUI) called the myCBR Workbench.

The myCBR SDK is written in the Java programming language and is the core of myCBR, containing all of its functionality. The SDK allows developers to use or extend the functionality of myCBR in custom applications. The current version (3.1) of the myCBR SDK only supports the similarity based retrieval step of the CBR-cycle.

The myCBR Workbench provides a GUI for implementing and modeling most of the functionality in the SDK. The GUI lets developers prototype and utilizes the SDK without having an in-depth knowledge of its inner functions. The myCBR Workbench supports importing predefined cases with the Comma-separated values (CSV) file format. It also includes a set of default similarity measures, which are visualized with graphs and tables, making modeling of concepts easier.

[1] myCBR: `http://www.mycbr-project.net/`

3.3 Case-based Reasoning Recommender Systems

Recommender systems mainly use either a collaborative or case-based approach to give recommendations [11]. The collaborative approach uses the rating and opinions of other users who have experienced a product or gone through the same process. The case-based approach, on the other hand, uses a case-base as the pool of information to pick recommendations from.

A case-based reasoning recommender system (CBR-RS) share many similarities with case-based reasoning systems (CBRS). The most noteworthy is that both types of systems have a case-base which they use to determine the best solution or recommendation. While a CBRS receives a query structured as a problem, a CBR-RS, on the other hand, receives a query structured more like a wish [30]. This wish includes user preferences and, if present, information from a user profile. The main difference between the two systems, however, lies in the way the systems propose a solution to a query. A CBR-RS usually gives several recommendations relevant to the wish of the user, while a CBRS typically gives a single solution which matches the query best. Similar to a CBRS, a CBR-RS use CBR as the underlying methodology. The use of similarity-based retrieval is a beneficial feature for both CBR-RS and CBRS, giving advantages compared to more traditional exact matching techniques such as conventional database retrieval and classical constraint satisfaction techniques [11].

3.4 Evaluating Recommender Systems

Recommendation systems are often evaluated and ranked based on their prediction power [32]. However, what is considered a relevant recommendation in a given scenario is often subjective. Therefore, it is challenging to evaluate whether a recommendation is satisfactory. For example, some users are interested in receiving recommendations which they are familiar with, while others may want recommendations that are more diverse to discover new items. Shani and Gunawardana [32] describes three levels of experiments to evaluate recommendation systems: offline experiments, user studies and online evaluation.

Offline Experiments When evaluating recommender systems with offline experiments, there are no actual users. Instead, a collection of data based on earlier user choices is used to simulate real user queries. The results of the queries can be used to evaluate the prediction power of the recommender. The queries should, to produce proper results, be similar to the queries that users will make in a production version of the system. Since an offline experiment requires no user interaction, it is a cheap and efficient approach of evaluating a recommendation system. It does however only evaluate a small set of possible user queries, and does not provide any information on user behavior of the system. An offline experiment can be used on different recommendation algorithms to evaluate them against each other. The data sets used to simulate user queries can either be designed by hand or by using a random selector on stored queries. Different types of biases can be introduced when generating the user queries, but methods such as reweighting or resampling can lower the influence of these biases.

User Studies In a user study the recommender system is tested on real users who are given tasks to complete. Their behavior and their satisfaction with the recommendations are then measured, and the feedback used to improve the system. User studies can use questionnaires to get feedback from users on their experience. The main advantages of a user study are that a broad range of questions can be answered and that a large amount of data, both qualitative and quantitative can be collected. The test users should represent the actual population of desired users, or else the results might be more vulnerable to biases such as non-response bias.

Online Evaluation By evaluating a recommender system with the online evaluation method, data on real behavior and user interaction can be collected from a live system. Several measures can be used to evaluate the system in an online evaluation, e.g., using different recommendation algorithms and measuring their use. To ensure a fair test of the different algorithms it is important that the users are redirected to each algorithm randomly. When the focus is only on evaluating the user interaction and design of the system, it might be more desirable to use only one algorithm and reduce the users' disorientation.

Chapter 4

Research Method

The purpose of this project is to create a prototype to improve students' motivation for applying for an exchange program by recommending courses and universities. The research questions were first created after defining the project's purpose and refined after identifying the background, motivation, relevant theory and related work through the preliminary research. The research questions were answered by evaluating a prototype created using the design and creation strategy. The final evaluation was done by using primarily quantitative data analysis on data generated by questionnaires and an offline experiment.

Oates'[26] model of the research process was used as a guideline for selecting the most appropriate research methods. The model and the specific methods chosen is displayed in Figure 4.1. This chapter first presents the chosen research strategy. Next, the data generation methods and their respective data analysis methods are described. Lastly, the practical and ethical issues which influenced and limited the research project are discussed.

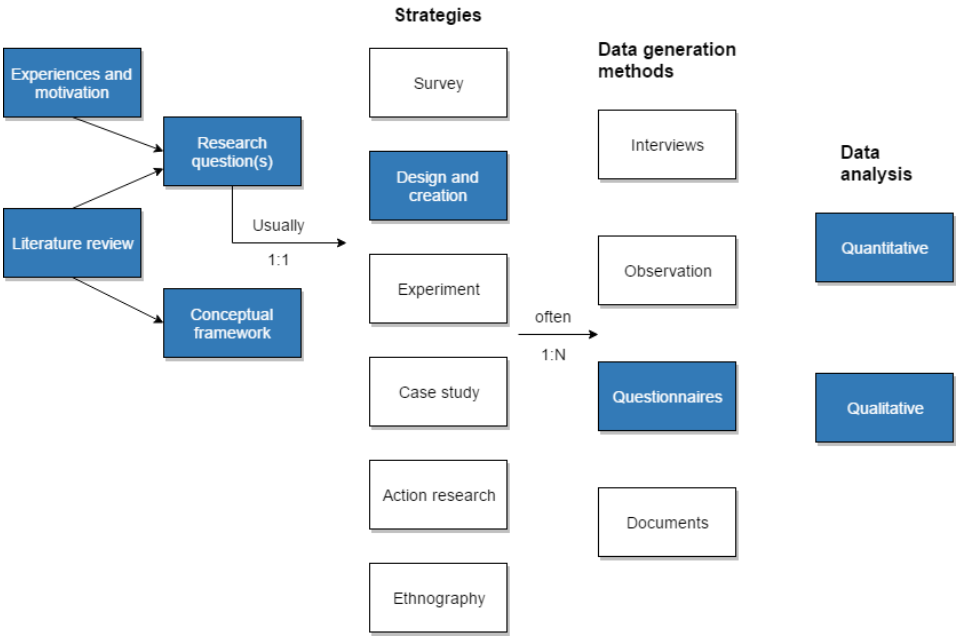


Figure 4.1: Outlined research process for the project

4.1 Research Strategy: Design and Creation

The design and creation strategy focuses on creating computer-based artefacts which in some way contributes with new knowledge [26]. An artefact can be either be a contribution to new knowledge in itself, a tool used to answer questions, or the end-product of a development process focused project. This strategy adds research properties to regular software development by including academic qualities such as analysis, explanation, argument, justification and critical evaluation.

This project utilized the design and creation strategy to develop a prototype named Utsida. Utsida is an information system (IS) built from the identified system requirements (sec. 2.4.1) and is used as a tool to answer both RQ1 and RQ2. The design and creation strategy follows the steps of the Design Science Research process model defined by Vaishnavi and Kuechler [36]. These steps are: Awareness, Suggestion, Development, Evaluation, and Conclusion.

4.1.1 Awareness and Suggestion

The awareness step involves finding and defining the problem domain, while in the suggestion step, the problem is refined into a tentative idea on how it might be solved. The domain was defined by the preliminary research (Ch. 2), through the production of research questions (sec. 1.3) and by gaining knowledge on relevant theory (Ch. 3). When the problem had been properly identified, a initial suggestion for the system was made. This suggestion included the first design details for the system which initiated the iterative development process.

4.1.2 Development

In the development step the idea formed in the suggestion step is implemented and a choice is made on what software development process to use. To answer the research questions it was necessary to produce data on user interaction and satisfaction. It was, therefore, essential that the web application part of Utsida was easy and intuitive to use. To ensure high user satisfaction, the software development of Utsida was done in an iterative process. This included communicating with potential users throughout the development and conducting usability tests. The spiral model, introduced by Boehm [7], is a common model for iterative development, and was chosen as the foundation for the software development process of Utsida. Each spiral in the model includes steps for analyzing risks and requirements, development and testing of a prototype, planning the next iteration, and determining the objective of the next iteration. Figure 4.2 shows the chosen development process of Utsida. The process was inspired by the spiral model, but has less risk assessment, validation and planning. The changes were done to implement and test new functionality more rapidly. After each development step, the prototype was tested on students through usability tests. In the next step, verification and validation, the results of the usability tests were used to evaluate the prototype based on the requirements and goals. The development process was then continued if needed or terminated if the prototype was deemed ready for the final evaluation.

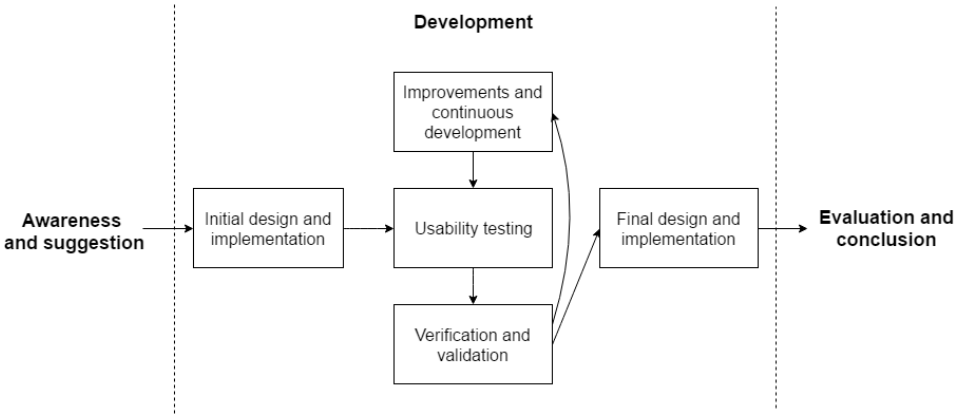


Figure 4.2: Iterative development process of Utsida

Usability Testing

The usability tests in each development cycle were performed on the web application to evaluate and ensure sufficient usability. The test participants consisted in a large degree of students who had an understanding of the current approach for applying for an exchange program. Students from different study programs were tested to ensure a broad specter of backgrounds. These tests helped find and resolve flaws in both the general functionality and design so that the system was as reliable as possible before conducting the final user study. Data was collected by interviewing the testers, observing their use of the system, and have them fill out a System Usability Scale (SUS)-schema [12]. During the interview the interviewee was asked questions such as "How would you expect this functionality to work?", "Was it easy to navigate the system?" and "Was it intuitive to understand how that functionality works?". After each test, the data was analyzed by using the SUS score [12].

4.1.3 Evaluation and Conclusion

In the evaluation step the developed prototype is assessed. The results and knowledge gained in the evaluation are then analyzed and discussed in the conclusion step. The final evaluation of Utsida and conclusion of the project was done by analyzing and discussing

the data generated by the chosen data generation methods. Several methods were reviewed. Because there was no client, or persons with high stakes in this project, interviews were not viewed upon as a feasible method to generate data. Data generation through documents was also not a viable method, because of no existing documents about the problem context. The observation method was used in the usability tests but was not relevant for the final evaluation. However, questionnaires and experiments were suitable ways to generate the data needed the answer the project's research questions and were, therefore, chosen as the methods to use. Using questionnaires made it possible to obtain generalized opinions on the system's motivational effect, and an experiment made it possible to objectively evaluate the suitability of the recommendations without user bias. Two questionnaires, questionnaire 1 and 2, and one experiment was used in the final evaluation. Questionnaire 1 was a study on motivational factors for going on exchange. While questionnaire 2 and the experiment were designed after known methods of evaluating recommender systems, namely the user study and offline experiment methods presented by Shani and Gunawardana [32]. The results of these data generations methods are presented in Chapter 6, and discussed in Chapter 7.

4.2 Questionnaires

Questionnaire 1 (sec. 4.2.3) was conducted during the development of the system with the goal of learning what students think are the most important motivational factors for choosing an exchange location. Questionnaire 2 (sec. 4.2.4) was conducted after the prototype was deemed ready and was a final evaluation and user study of the prototype, Utsida.

Both questionnaires were self-administrated, which promoted the opportunity to collect data from a representative number of students at once. The target group of the questionnaires were students that had some experience with exchange studies. To acquire a large amount of students who were in this group, a cooperation with the OIR was made, and they agreed to publish the questionnaires on their Facebook page with approximately 3300 followers, most of them being in the target group of this research project.

In order to keep the margin of error to a minimum it was important to get feedback from as many students as possible. The target population of the questionnaires was students that have been on exchange from NTNU, and students that were interested in exchange. The population was roughly estimated to be between 10000 to 14000 students. According to Krejcie and Morgan [18] a sample size of approximately 370 is optimal (i.e. to give a margin of error less then 5%) with a population size of 10 000. However, considering the limitations of budget, time and marketing resources, the expected sample size in this study was considerably lower than the optimal size. Hence, a sample size goal of 100 students was set for the questionnaires. With a population of 10 000 and confidence level of 95%, this would give a margin of error of approximately 10% [39].

The majority of the questions in the questionnaires were closed questions formed as either: a Likert Scale [2], a semantic differential scale [28] or with "Yes", "No", or "Don't know" options. Also, the students were asked optional open questions where they could articulate their opinions freely.

4.2.1 Validity and Reliability

Content validity is concerned with whether the questions in the questionnaire represents a good sample of the actual problem to be investigated [26]. This means that the selected questions should cover all the different evaluations needed to answer the research questions accurately. Feedback from potential participants and pilot tests were used to evaluate the validity, and the questions were formed with the possible evaluation criteria of the research questions in mind. For RQ1 the criteria is the use of the system and its motivational effect, while for RQ2, the criteria is if the system recommends viable universities and courses.

Construct validity means that the questionnaire is evaluating what it was intended to evaluate, and not other aspects [26]. Both questionnaires were designed to ensure the construct validity by using clear and straightforward questions that were in the scope of the research questions. The questionnaires were designed in an iterative process to remove or change possible irrelevant or poorly formulated questions.

The measurement of reliability finds out whether the questionnaire yields consistent results. The reliability is difficult to assess as the questionnaires were only published once. Evaluating several times with part of the participants could also have distorted the data by giving them more time to be familiar with the prototype. However, one way to measure the reliability or internal consistency is using Cronbach's alpha [6], which was done on the Likert scale questions in questionnaire 2.

4.2.2 Bias

When using questionnaires as the primary method to collect data, several risks of bias can be introduced, including central tendency bias, acquiescence bias, social desirability bias and non-response bias [16]. This section presents these biases, how they could affect the results and how the risk of bias was reduced.

Central tendency bias is a risk when the questions have extreme response categories and is especially relevant for the Likert scale and Semantic Differential Scale. This bias means a participant is more likely answer in the middle of the scale. A lower point scale was used in questionnaire 2 to reduce the tendency of answers being in the middle of the scale. Both questionnaires also aimed to have clear questions that reduced the likelihood of participants to answer in the middle of the scale.

Acquiescence bias is when the participant is likely to answer positively on a question [13]. To reduce the acquiescence bias, the questions were designed to not be very positively based. However, due to the nature of the questions having less social desirability and not being of a sensitive type, an acquiescence bias was not likely to have a large impact.

Social desirability bias influence a participant to deny undesirable traits or positively answer questions that are socially desirable. Questions were striven to be neutral statements where the participants could state their degree of agreement. An example is the use

of the Likert-type question format.

Non-response bias is when respondents' answers are considerably different from the possible answers of those that do not respond. Considering the response on the questionnaires was a small proportion of the full population, non-response bias was unavoidable. Those who did not participate might think the research sounds uninteresting, and if they participated, they might respond with more negative answers. Oppositely, the target population for the questionnaires were students who already had an interest in exchange studies, which might influence responses to be more positive.

4.2.3 Questionnaire 1: Motivational Factors

The goal of questionnaire 1 was to gain knowledge on important motivational factors for students when they choose an exchange location and university. The results from this questionnaire was also used to assign the weights of the attributes in the exchange experience concept, and used to justify what attributes to include.

The main question in the questionnaire was closed and formed according to a seven-point semantic differential scale, shown in Figure 4.3. A seven-point scale was chosen instead of the more standard five-point scale to better fit the model in the myCBR Workbench where the weights are represented in a range from 1-10. The participants ranked each motivational factor according to their opinion of the factor's decisiveness. Questionnaire 1 also included an open question which asked for input on other important motivational factors. See appendix A.2 for all the questions included in questionnaire 1.

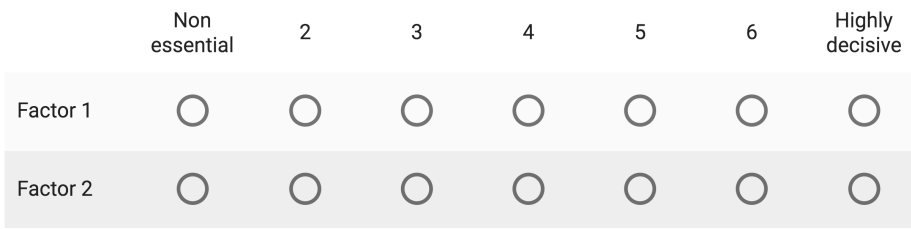


Figure 4.3: Question format in questionnaire 1

Data analysis

The numerical data from the questionnaire was analyzed in a quantitative manner, by calculating the mean, standard deviation and coefficient of variation for each attribute. The coefficient of variation (i.e. the spread relative to the mean of the data) was calculated to rank each attribute against each other. The resulting rank was then used to choose the weights of the attributes in the CBR-RS.

The open question was analyzed in a qualitative manner by using theme analysis [26]. Each answer was first reduced to only contain important sections related to the goal of the question. These important sections were then reduced to one or several common themes, and, finally, the number of occurrences of each theme was counted and ranked by importance.

4.2.4 Questionnaire 2: User Testing of Utsida

The three goals of questionnaire 2 was to: 1) evaluate and analyze how Utsida affects students' motivation for applying for an exchange program; 2) find out whether students who have already been on an exchange program thinks Utsida could have made their process easier; and 3) evaluate if Utsida gave relevant recommendations to the user.

The questionnaire was structured in two parts; one for reviewing the motivation and use of the application, and one for evaluating the CBR-RS' recommendations. All the questions from both parts can viewed in Appendix B.3.

Part 1: Motivational Effect and Use

The goal of part 1 was to find out what effect Utsida may have on students' motivation to apply for an exchange program (RQ1). The participating students got slightly different questions depending on whether they had participated in an exchange program or not. The questions followed the Likert-type format [20] with a five-point scale where each option represents the degree of agreement to a given statement, shown in Figure 4.4

Figure 4.4: Question format for selecting level of agreement on a statement in questionnaire 2

Part 2: Evaluation of Recommendations

The second part of the questionnaire targeted RQ2 by letting participants evaluate the recommendations they received from Utsida, and evaluate the recommendations from two pre-defined queries. The question form was designed to have simple "Yes", "No", "Don't know" answers to easier determine the amount of participants that received relevant recommendations. The pre-defined queries (App. B.1), were included to evaluate the recommendations for a specific query and to avoid the evaluations being affected by a participants' preference for a query. The evaluation of the pre-defined queries included a question which asked the participating students to rate the recommendations on a scale of suitability (Fig. 4.5).

Figure 4.5: Question format for evaluating suitability of recommendations for a pre-defined query in questionnaire 2

Data Analysis

The data analysis on the results from questionnaire 2 was done in a similar fashion as in questionnaire 1 by using quantitative analysis on the closed questions and simple qualitative analysis on the open question. The closed Likert-type questions produced ordinal data, meaning that the standard range between the values on the scale can not be deter-

mined. The data was, therefore, analyzed by using frequencies and mode. For the “Yes”, “No” or “Don’t know” questions however, only percentages and number of answers for each option is used. Cronbach’s alpha was used to asses the internal consistency of the Likert scale questions regarding motivation.

4.3 Offline Experiment on the CBR-RS

An offline experiment method introduced by Shani and Gunawardana [32] was used to evaluate the relevancy of the recommendations of the CBR-RS, and thus contributed to answering RQ2. The experiment is a specific method to evaluate recommendation systems in an offline setting without user influence, and should not be misinterpreted as the *Experiment* research strategy.

The offline experiment was conducted with 20 queries (App. C.1) that were generated to simulate 20 unique users. Each query was sent to two different concept configurations in the CBR-RS. One with adjusted similarity measures, taxonomies and weights, and one simulating standard exact match search. The simulation was done by only using exact match similarity measures and equal attribute weights. The scores for each recommendation was then given according to the score metric in Table 4.1, and the total score for each configuration was the mean score of each query. The hypothesis was that the configuration with adjusted similarity measures and weights would score higher than the configuration with a standard exact match search.

This method made it possible to test the CBR-RS with a large set of inputs at a low cost. One downside of the method is that the queries only represent a small part of the possible user searches. This was mitigated by selecting a wide range of possible attributes and users profiles. Another disadvantage was that the offline experiment did not evaluate the CBR-RS with actual users, but it was mitigated by conducting continuous usability tests in addition to the user study with questionnaire 2. The Course attribute is not included in this experiment because the relevancy of the attribute is difficult to objectively measure and score without user feedback.

Rating scale for recommendations (0-10 points)			
Attribute	0 points	1 point	2 points
Department	wrong faculty, wrong department	correct faculty, wrong department	correct department
Year	before 2011	2011-2013	2014-2017
Geographic Location	wrong country, wrong continent	correct continent, wrong country	correct country
University	no match		perfect match
Language	no match	the language is part of the list	perfect match
Ratings	0.5 points for each rating in range (-1, rating +1)		

Table 4.1: Score matrix used to evaluate the recommendations in the offline experiment

Data Analysis

The offline experiment generated quantitative data that was statistically analyzed. The central tendency was analyzed with the mean and standard deviation measures, and a paired t-test was conducted to define the significance of the result. The confidence level of the t-test was 95% with a p-value limit of 0.05.

4.4 Ethical Issues

Some ethical issues regarding data collection and storage were considered in the research project. The designated data collection methods kept all participants anonymous, but some meta-data was collected, such as which faculty and university the participants belonged to. Furthermore, as an incentive for students to answer the questionnaires, lottery prizes were promoted for participating. To be able to enter the lottery, the participants had to enter their e-mail. The lottery was entirely optional and participants could register their answers without entering the lottery. The answers were stored in a Google Spreadsheet anonymously, and the entered e-mails were scrambled in a random order so that they were not linked to their answers. This way, there was no way to identify each answer. After the lottery was completed all the emails were deleted and only the winner contacted. Informed consent was maintained in the questionnaires and usability tests by informing

the participants about the purpose and goals, and by stating that the entire process for all participants is entirely voluntarily and that they can stop any time they want.

During the user study of the system, the users created an account which included storing their chosen password, e-mail, name and department at NTNU. This information could not be linked to the questionnaire they took, but it still raised an ethical issue, considering the authors had access to this data. During the development of the system, steps were taken to ensure the confidentiality of the users of the system by using encryption and proper security measures.

4.5 Practical Issues

This research project was performed by two students, with one supervisor to guide the project. There was a minuscule amount of monetary funds, and the available project duration was limited to nine months. The data which served as the results was primarily produced by real users. With the minimal budget, a practical issue was to produce incentive for enough students to answer the surveys. Furthermore, because of the limited time, both larger administrated tests and self-administrated tests were considered out of reach.

Chapter 5

Implementation

To be able to answer the research questions using the methods described in the previous chapter, a prototype needed to be developed. This chapter covers the implementation and architecture of that prototype, named Utsida, and introduces its different parts, components, design and usability choices, and test environment.

5.1 System Overview

Utsida consists of two sub-systems, where the first part is a web application, and the second part is a case-based reasoning recommender system (CBR-RS). Each part primarily focuses on one of the research questions. The web application is designed to increase the motivation (RQ1), while the CBR-RS is designed to find out how suitable CBR is to give recommendations on exchange universities and courses (RQ2). The recommendations made by the CBR-RS could also affect the motivation (RQ1). Utsida was chosen as the name of the system due to being a counter opposite to NTNU's central system *Innsida*, and the meaning of the word *inside*. Utsida gives a relation to something on the outside, in this case going on an exchange program. The design and implementation of Utsida was based on the system requirements (sec. 2.4.1) identified in the preliminary research.

As illustrated by Figure 5.1, the web application part communicates with the CBR-RS part. Recommendation queries are sent from the web application by a user and received by the CBR-RS through its Representational State Transfer (REST) Application Programming Interface (API). These queries are in turn matched against the entire case-base in the CBR-RS. Finally, all the cases are given a similarity score based on the user query and the most similar cases are returned to the web application. The user can then select a specific case from the several recommendations shown in the web application.

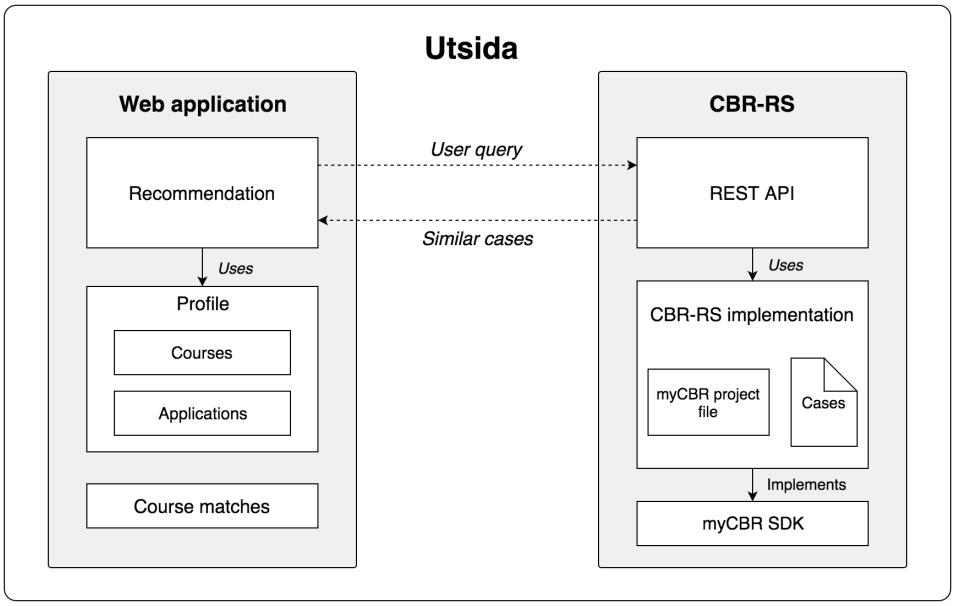


Figure 5.1: Architectural overview of Utsida

5.2 Data and Information

Utsida requires up to date organizational data about NTNU. This includes all the courses, faculties, and departments at NTNU. The required organizational data was gathered with the Faculty of Information Technology and Electrical Engineering organizational (IE)'s API[1]. This API was, however, recently discontinued, so to update this data after 2016, a

[1]NTNU organizational API: `http://www.ime.ntnu.no/api/`

new data source would have to be found. Furthermore, the course matches component (sec. 5.3.1) in the web application relies on information on previously approved course matches[2]. To avoid a cold start, some approved course matches from IDI were parsed and used in the component. The most critical data, however, was the data used to create cases for the CBR-RS, namely the *experience reports*. All data used in Utsida was analyzed according to the six primary dimensions for data quality assessment [3] to gain awareness of the possible limitations before deciding on Utsida's database models.

5.2.1 The Experience Reports

Most students who go on an exchange program from NTNU have to write obligatory reports about their exchange experience to the OIR. These reports are available to the public, and are written in the period between 1999 and 2016. Because the reports include a large amount of information on exchange experiences, they were used as the data foundation for the cases. Typically, the case-base in a CBR-RS start off with a low amount of cases and gradually expand as new cases are retained. In this project, however, it was essential to have an initially large number of cases. Therefore, all of the current public experience reports were parsed and used as cases in the CBR-RS' case-base.

5.2.2 Parsing the Experience Reports

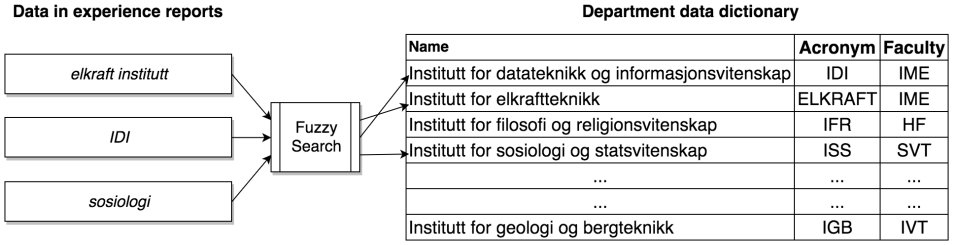
The experience reports are essentially large text files in HyperText Markup Language (HTML) format. To be able to use them in the CBR-RS, all the useful information in the files had to be extracted and parsed to cases in a CSV file, which is the file format used for case-bases in myCBR. This was done with a custom script written in the Python programming language. In short, all available experience reports were first downloaded as HTML files and stored in a directory. Next, the script looped through all the HTML files, read the internal data, and finally stored all the chosen data in a CSV file. This way, each

[2]Course matches: Sets of courses where one or more courses from another university are approved by advisers as a replacement for a course at NTNU

row in the CSV file would correspond to one case with the attributes given by Table 5.1. This process resulted in a CSV file with 8702 cases which formed the case-base.

A large part of the text fields in the experience reports were free text, which means that the format of each field was decided by the author and did follow any standard. The parsing script, therefore, had to use specific methods to parse each attribute to the correct format. Two methods were primarily used to parse the data; data dictionary look-ups and approximate string matching, also called fuzzy searching. Edge cases that were not detected by these techniques were solved with basic string operations. In the end, most of the data in the experience reports was transformed into usable data in the CBR-RS. It was, however, not possible to ensure that all of the data in the experience reports were parsed to a correct format while still keeping its original meaning.

Fuzzy Searching is an algorithm used to determine the similarity between two textual inputs. It calculates a similarity score between two textual inputs by recognizing patterns and similarities in the two inputs. By using fuzzy searching in the parser script, the data in the experience reports could be mapped against a predefined set of attributes, and return the attribute with the highest similarity. For example, the department field in the experience reports had a significant variation of abbreviations and ways of spelling. By creating a dictionary with all departments and faculties, fuzzy searching could be used to find the department in the dictionary with the highest similarity to the department in the experience report. This is illustrated in Figure 5.2. This method was also used on the university field to avoid creating multiple instances in the database. Several more Python dictionaries were configured to be able to map the necessary data in the experience reports to a correct format. These includes dictionaries for countries, continents, languages, universities, departments, and faculties.



Name	Acronym	Faculty
Institutt for datateknikk og informasjonsvitenskap	IDI	IME
Institutt for elkrafteknikk	ELKRAFT	IME
Institutt for filosofi og religionsvitenskap	IFR	HF
Institutt for sosiologi og statsvitenskap	ISS	SVT
...	...	...
...	...	...
Institutt for geologi og bergteknikk	IGB	IVT

Figure 5.2: How fuzzy searching is used to convert fields in the experience reports to a standard format defined in a dictionary.

5.3 The Web Application

This section presents the different components of the web application, the tools and frameworks used, and reasoning for design and usability choices.

5.3.1 Components

The web application part of Utsida is divided into three main components, as Figure 5.1 depicts: *Recommendation, Profile* and *Course matches*. The aim of the components as a whole is to answer RQ1 and increase the motivation of students to go on an exchange program. Each component targets one or several of the requirements of Utsida, listed in Table 2.1 The requirements are referenced in a REQ.# format, with # as the requirement number. REQ.5 is handled by the web application home page and not by any component.

Recommendation

The recommendation component targets REQ.1. This component handles all the communication with the CBR-RS part; Queries made by users are sent in the JavaScript Object Notation (JSON) format to the CBR-RS through HyperText Transfer Protocol (HTTP) requests, which returns all of its cases and their respective similarity score, also in JSON format. In this component, a user can select their preferences and requirements for an exchange program. The user is then presented with a list of the best suiting universities

for their query (Fig. 5.3). The universities in the list can each be selected to show the best matching experiences for that university (Fig. 5.4). Each experience contains a list of courses taken and serves as a recommendation.

Course Matches

The course matches component target REQ.2, and its purpose is to display all approved course matches stored in the web application. The component filters course matches by which university they are approved for, and supports adding, deleting, commenting and editing of course matches by student advisers. For a student, this component serves as a resource where they can: find replacements for their courses at NTNU, get information about the courses and the approval of them, and save the course match to their profile.

Profile

The profile component handles all profile functionality, such as authentication, and contains the required personal information about each student. One especially important information is the user's department, which is stored in the profile and used by the web application in a recommendation query. The component has two sub-components: Courses and Applications.

Courses is a sub-component of the profile component and targets REQ.3. It contains functionalities which enable a student to organize their application process. This includes storing courses at home, courses abroad and course matches. From the course page, shown in Figure 5.5, students can also manually match abroad- and home courses to create new course matches for their applications. This component replaces a large part of the course match approval process and is essential for the digitalization of the process.

Applications is another sub-component of the profile component and targets REQ.4. It stores user submitted applications that contains the course matches a user wants to get approved. It also maintains the functionality needed to create applications. For students

and student advisers, the application component serves as a digital version of the current paper-based application process. In this component, advisers can approve, reject, and remove applications, while students can submit and track the status of their application. This component is an extension needed to complete Utsida as a possible replacement of NTNU's current approach of approving courses for an exchange program.

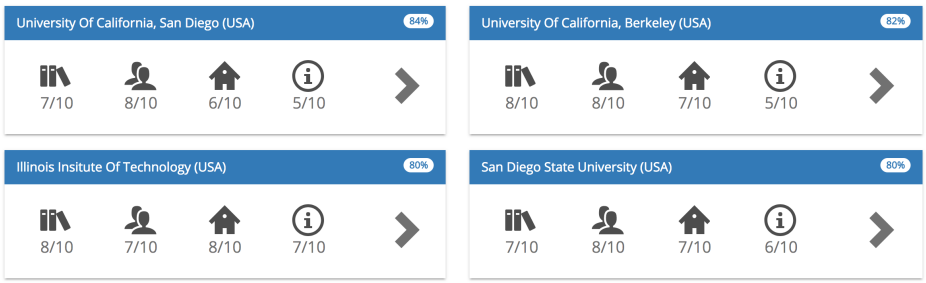


Figure 5.3: Recommendation view

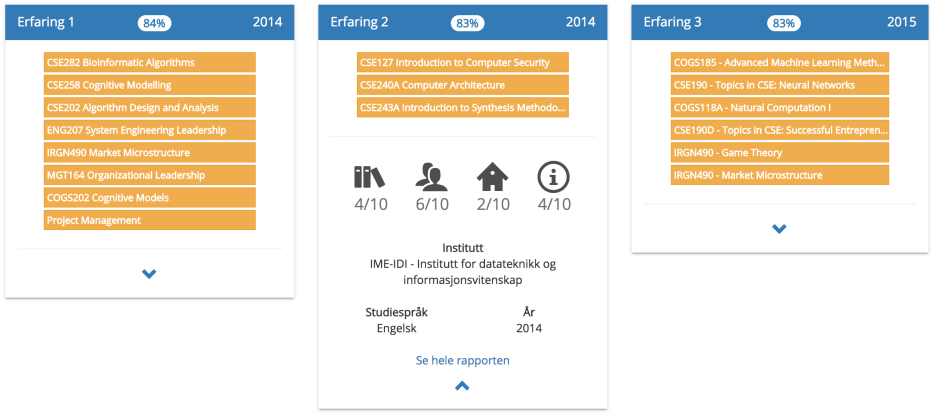


Figure 5.4: Course recommendation view after selecting university

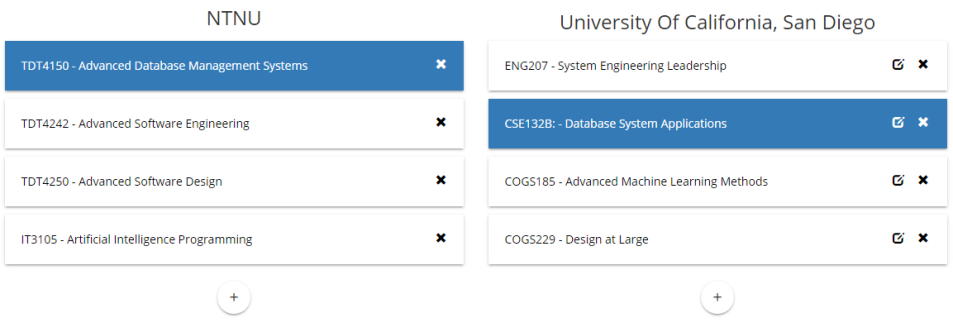


Figure 5.5: View for saving, viewing and matching courses

5.3.2 Tools and Frameworks

When developing the web application part of Utsida, several tools and packages were used to simplify the development process and to implement functionality that was not feasible to develop from scratch.

To enforce a responsive and aesthetic design, Bootstrap 3[3] was used as the front-end Cascading Style Sheets (CSS) framework. jQuery[4] was used to simplify HTTP requests, which this system largely relies on both internally in the web application and between the web application and the CBR-RS. A collection of packages for the Python language was also used. These packages include FuzzyWuzzy[5]: a packages which handles apprximate string matching, Requests[6]: a package for writing HTTP requests from Python, and Python-social-auth[7] combined with Dataporten-auth[8] to enable the use of authentication methods such as Feide, the authentication service used at NTNU.

Django[9], a web framework using the Python programming language, was chosen as the web framework to use in the web application. Django is maintained by the Django Software Foundation (DSF). The DSF calls it an MTV-framework, meaning *Model, Template,*

[3]Bootstrap: `http://getbootstrap.com/`
[4]jQuery: `https://jquery.com/`
[5]FuzzyWuzzy: `https://pypi.python.org/pypi/fuzzywuzzy`
[6]Requests: `http://docs.python-requests.org/en/master/`
[7]Python Social Auth: `http://python-social-auth-docs.readthedocs.io/`
[8]Dataporten-auth: `https://pypi.python.org/pypi/dataporten-auth/0.1`
[9]Django: `https://docs.djangoproject.com/en/1.10/`

View, which in practice functions as a Model View Controller (MVC) pattern. Django includes many built-in features: an SQLite database, a predefined directory structure, controllers to handle communication between the different components, and many others. Another important feature is Django's built-in security packages, which handles security risks such as Cross-Site Request Forgery, authorization, and Cross-Site Scripting. Django also includes an administrator panel, which enables management of database models with a simple interface. In conclusion, the reason for choosing Django as the main framework was because it offered a complete package that reduced the time used on structural difficulties and made it possible to focus on the usability and evaluation methods.

5.3.3 Design Choices

The design in terms of both functionality and aesthetics was an integral part of the development of the web application part of Utsida. Both of the questionnaires in the project were conducted remotely and were self-administered, which removes the option to help the participants with possible issues. Therefore, Utsida had to be designed in a way which would produce as little issues and hardships for the users as possible.

The design process of Utsida is inspired by the User Centered Design process, introduced by Norman and Draper [24]. Potential users of the system were included throughout the development process, either with small interviews or larger usability tests. The results of these tests resulted in several identified problem areas. Measures were implemented to reduce the problem areas and strive for a user-centered design. Some of these measures were:

- Personalized items: Buttons and links to parts of the application that contain user specific data are prefixed with the word "*My*", for example: *My courses, My profile, My applications*.
- Inline help text: Some parts of the application contain help text for users that may need it. The help texts are shown by hovering or clicking on a question mark button.
- Streamlined process: The typical chronological use of Utsida is displayed as a nu-

merated process to guide the user, as shown in Figure 5.6.

Figure 5.6: Home page of Utsida. Showing the process flow help text.

5.4 The CBR-RS

This section describes how the CBR-RS was modeled with the myCBR Workbench and how it was implemented to give the most relevant recommendations. Furthermore, this section includes details on the case representation, the similarity measures used, how the weighting of the attributes in the exchange experience concept was done, and how the system itself is structured.

5.4.1 Overview

The CBR-RS is a customized version of an open-source Java Spring[10] application that implements the myCBR SDK. To serve requests from a web application, the Java application uses a REST API implemented with the Spring framework. This architecture made it possible to let the web application and the CBR-RS be two independent systems, only communicating through HTTP requests. The CBR-RS imports a custom myCBR project file which contains the exchange experience concept's case representation, the similarity

[10]Spring: `https://spring.io/`

measures used for each attribute and the attributes' weight values. It also imports a CSV file containing all the cases, used as the case-base.

When a request is made by a user in the web application, a query is first sent to the CBR-RS. When the CBR-RS receives the query, it performs the first step of the CBR-cycle by using the retrieve operation in the myCBR SDK. The myCBR SDK then returns the exchange experiences (i.e. cases) with their respective similarity scores. The exchange experiences are further ordered from highest to lowest similarity, and finally, the exchange experiences with the highest similarity scores are returned to the web application. Figure 5.7 displays how a request made by a user in the web application traverse between the two systems to return a list of the most similar exchange experiences.

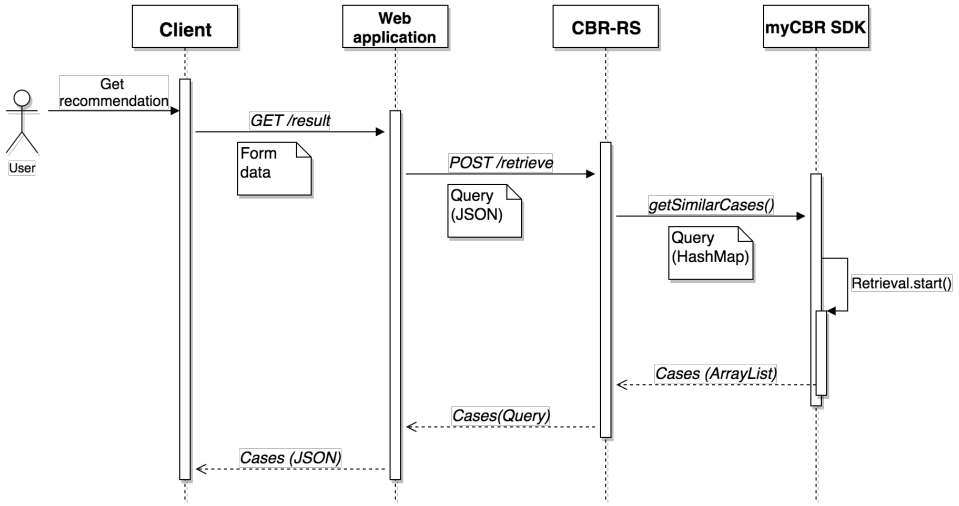


Figure 5.7: Sequence diagram for a recommendation request in Utsida

5.4.2 Case Representation

The exchange experience reports are the underlying data the CBR-RS use to give recommendations. These reports contain a significant amount of information and not all of it is relevant. Only a small part of the data was used to ensure that Utsida gives relevant recommendations and is straightforward to use. The first choice of information to use for the initial case representation were based on personal experiences and a study by Cubillo et.

al. [21] that identified groups of important motivational factors. This representation was, however, extended in an iterative process of usability tests and then finalized by using the results of questionnaire 1 (sec. 4.2.3). The final case representation is displayed in Table 5.1.

Table 5.1: Case representation in the CBR-RS

Attributes	**Example case values**
Department	IME-IDI - Institutt for datateknikk og informasjonsvitenskap
Continent	North America
Country	USA
University	UCLA
Language	English
Study Period	2012
Academic Quality Rating	8
Social Quality Rating	4
Ease to find- and quality of residential	5
Support and reception at university	6
Subjects Taken	COMP1927 - Computing 2 MATH3220 CS4210 DATA101

The Problem Part

The problem part of a case is all attributes which denote a user's preferences for a query and profile information about them. The following list elaborate on the different attributes used in a case.

- **Department:** A profile attribute which details the department a user belongs to. It is considered the most important attribute, as it highly indicates what kind of courses (i.e. solution) that are relevant to a user.
- **Continent:** A preference attribute that implies what continent a user wish to travel to. Allows for a less precise location than a country.
- **Country:** A preference attribute that implies what country a user wish to travel to.

Serves as the most precise geographical location a user can select.

- **University:** A preference attribute that lets a user include a specific university they wish to go to.

- **Language:** A preference attribute that represents the language the user wish to study in.

- **Study Period:** A hidden attribute that is the current year for a problem. In a case it is the year an exchange experience was written. The attribute is included to make older cases less relevant to the user, as university courses often change over time.

- **Academic Quality Rating:** A preference attribute that is a conjunction of an experience report's *Academic Quality*, *Special Competence Gained* and general *Quality of Academic Opportunities*. Each of these aspects is related to the academic quality of an experience, and are weighted differently.

- **Social Quality Rating:** A preference attribute that is a conjunction of an experience report's rating of *Social Quality*, *Leisure Activities*, *Girlfriend/friends*, the social interaction with *Students from the University*, *Other Foreign Students* and *Norwegian Students*. Each of these aspects is related to the social quality of an experience, and are weighted differently.

- **Ease to find- and Quality of Residential:** A preference attribute that is a conjunction of an experience report's rating of *Residential Dissemination* and *Residential Quality*. Each of these aspects is related to the residential part of an experience, and are weighted differently.

- **Support and Reception at University:** A preference attribute that is a conjunction of an experience report's rating of *General Reception* and *Administrative Support*. Each of these aspects is related to the support a student received at the university in an experience, and are weighted differently.

The Solution Part

The solution to a query is given by the set of courses in similar experiences. The university is however also a part of the recommendation and is used to filter the experiences, and can, therefore, be viewed as part of the solution.

5.4.3 Similarity Measures

Each attribute in the exchange experience concept was assigned a designated similarity measure. These measures returns a similarity score based on the similarity between the attribute in a query and the coherent attribute in a case. Following, each of the used measures are detailed and explained.

Rating Similarity

In the exchange experience concept, the attributes *AcademicQuality*, *SocialQuality*, *ReceptionQuality* and *ResidentialQuality* are conjunctions of different ratings in the experience reports. Each rating ranges between 1-10 and reflects how the student experienced these categories. To calculate the similarity between a rating in a case, and a rating from a query, the CBR-RS uses Symmetric Difference Determined Similarity [5] with a linear function (equation 5.1). It describes the decrease in similarity at a linear rate.

$$f(d) = \begin{cases} 1: & d < min \\ \frac{max-d}{max-min}: & min \leq d \leq max \\ 0: & d > max \end{cases} \tag{5.1}$$

Continent Similarity

The continent attribute is a *symbol* type, meaning its allowed values are a pre-determined set. The similarity between continents is calculated with an $m\,x\,n$-matrix, where a mapping between two continents gives the similarity, as displayed in Table 5.2. The values

were chosen to represent similarity in culture, language, location, and general university structure.

Table 5.2: Matrix for calculating similarity between continents

Attribute value	South America	Asia	Europe	Africa	North America	Oceania
South America	1.0	0.0	0.2	0.0	0.3	0.0
Asia	0.0	1.0	0.0	0.0	0.0	0.0
Europe	0.2	0.0	1.0	0.0	0.2	0.2
Africa	0.0	0.0	0.0	1.0	0.0	0.0
North America	0.3	0.0	0.2	0.0	1.0	0.2
Oceania	0.0	0.0.0	0.2	0.0	0.2	1.0

Country and Department Similarity

The Country- and Department attributes are structured as a taxonomy. A taxonomy can be used as a type of similarity measure [30] and gives a similarity between similar groups of objects. For the country attribute, the taxonomy gives a similarity if the country resides within the same continent (i.e. group) as the country in the query. If the countries in the query and a case are the same, the similarity score will be 1.0. If the countries are different, but still resides within the same continent, the similarity score will be set according to Table 5.3. The same taxonomy is also in place for the Department attribute, however, the departments are instead grouped by faculty. The similarity score between departments in the same faculty is set to 0.3.

Table 5.3: Taxonomic similarity of each continent

Continent	**Similarity**
North America	0.2
South America	0.3
Oceania	0.4
Asia	0.3
Europe	0.2
Africa	0.4

Language Similarity

The language attribute is represented as a symbol type, which allows the attribute to contain multiple languages. If a case contains more than one language, the similarity measure will return the mean similarity to all languages in that case. If a case only contains one language, it will either return a similarity score of 1.0 (if the language in the query and case are is the same), or else 0.0. For example, if a query contains the language *Spanish* and a case contains the languages *English, Spanish, Portuguese*, the similarity score will be 0.33.

University Similarity

The university attribute is a string type. The similarity measure between two universities returns 1.0 if they are a complete match, or 0.0 if they differ. It would be desirable to have partial similarity for two strings, but this functionality is not included in myCBR.

Study Period Similarity

The similarity measure used for the study period attribute is defined by a declining graph, where the similarity score decrease according to the distance in years between the query and a case. Figure 5.8 illustrates the graph.

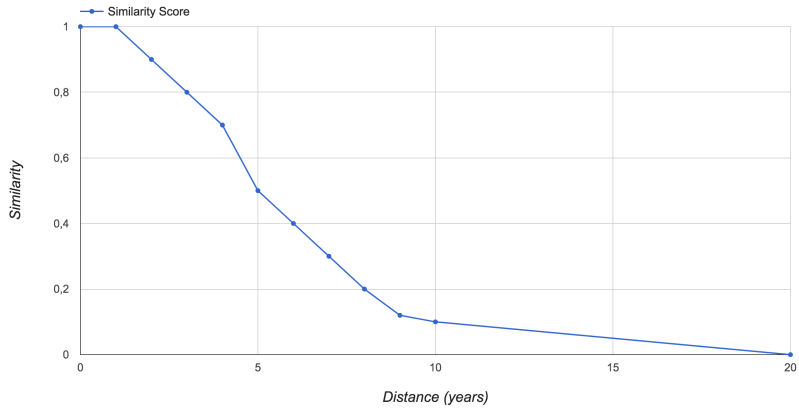


Figure 5.8: Study period similarity given by the distance in years between a query and a case.

5.4.4 Summation Function

Multi-Criteria Decision Making Methods (MCDMM) are used in myCBR as summation functions that combine the similarity score of each attribute in the concept, with regards to the attribute's weight. The myCBR Workbench has two MCDMMs, the Weighted Sum Model (WSM) and the Euclidean Distance. For single dimensional problems, which are the types of problems in the CBR-RS, the WSM can be used without difficulties, and is probably the most common approach [35] (equation 5.2 [15]).

$$A^*_{WSM-score} \quad = \quad max_i \sum_{j=1}^{n} a_{ij} w_j, \quad for \quad i = 1,\ 2,\ 3,\ ...,\ m. \tag{5.2}$$

When the CBR-RS receives a query, a decision matrix is calculated for each case. This is an $(m \times n)$ matrix where each row contains the attribute values and weight of a case, as shown in Table 5.4. The WSM is used on this matrix to give each case in the case-base a similarity score between 0.0-1.0.

Table 5.4: Decision matrix example for calculating similarity between a query and the cases in the case-base

	Attribute	Country	Language	...	...	AcademicQuality
	Weight	$w_{country}$	$w_{language}$	...	...	$w_{country}$
Cases	$case_1$	sim_{case_1}	sim_{case_1}	...	...	sim_{case_1}
	$case_2$	sim_{case_2}	sim_{case_2}	...	...	sim_{case_2}
	.	.	.	...	...	.
	.	.	.	...	...	.
	$case_m$	sim_{case_m}	sim_{case_m}	...	...	sim_{case_m}

5.4.5 Attribute Weights

Each attribute in a concept has a value and an associated weight. This is essentially a number denoting the importance of that attribute. The results of questionnaire 1 (Table 6.1), were used to decide the final weight of the different attributes in the exchange experience concept, shown in Table 5.5.

Table 5.5: Weighting of the concept's attributes

Attribute	Weight
Department	7.0
Continent	2.5
Country	4.0
University	4.0
Language	4.0
Study period	4.5
Academic Quality Rating	3.0
Social Quality Rating	3.0
Ease to find- and quality of residential	1.5
Support and reception at university	2.5

5.5 Test Environment

The test environment used by Utsida is an essential part of the project because it was the platform used for the user study in questionnaire 2. Utsida was hosted on a Ubuntu 16.04[11] virtual server provided by IDI with the domain name utsida.idi.ntnu.no. The specific HTTP web server used was Apache2[12], and it was configured with HTTPS to enable secure and encrypted user sessions. Support for registering and logging in with users university accounts was implemented to reduce the effort for users to participate in online usability tests and user study. The service platform "Dataporten"[13] provided by UNINETT made it possible to use the university account registration through UNINETT's federated authentication service (FEIDE).

Error reporting with email was configured on the server so that alerts were given in case of possible downtime or server errors. Error reporting made it possible to quickly resolve potential bugs or errors in the application and ensure high user satisfaction during the user tests. Specific user behavior and statistics was tracked by using Google Analytics[14], which was useful to increase the understanding of user behaviors. The most important Google Analytics data used in this project were number of users, average session time, device type and activity flow.

[11] Ubuntu: `https://www.ubuntu.com/`
[12] Apache2: `https://httpd.apache.org/`
[13] Dataporten: `https://www.uninett.no/en/dataporten`
[14] Google Analythics: https://analytics.google.com/

Chapter 6

Results

This chapter provides all final results yielded by the evaluations of the prototype, namely the questionnaires and the offline experiment on the CBR-RS.

6.1 Questionnaire 1: Motivational Factors

The result from questionnaire 1 gave valuable insights on students' motivational factors for exchange, and was also used to weight the different attributes in the CBR-RS (sec. 5.4.5). The questionnaire had 84 participants of which 52 of them had participated in an exchange program. 67 were students at NTNU and 17 were students at other higher educational institutions in Norway. All of the answers included numerical ratings between 1-7, describing how important the participant thought each of the attributes in the questionnaire was. Additionally, 36 of the replies included an optional textual answer. Table 6.1 shows the values for each attribute scaled to a range between 1-10, while Appendix A.1 displays the result of the theme analysis, that is, the most frequent category of factors in the textual answers.

Table 6.1: Questionnaire 1, motivational factors result. Ranked by CV. N=84.
*SD: Standard Deviation, CV: Coefficient of Variation

Attribute	Mean (1-10)	SD (1-10)	CV (0-1)
Language	8.2	1.87	0.23
Social quality	7.02	1.75	0.25
Academic quality	7.11	2.12	0.30
Administrative support and reception	6.5	2.27	0.35
Country/Continent	6.84	2.45	0.36
Quality residential	5.27	1.92	0.36
Cost of living	5.66	2.13	0.38
Climate/weather	6.22	2.53	0.41
Availability of residentials	5.31	2.22	0.42

6.2 Questionnaire 2: User Study of Utsida

Questionnaire 2 had 40 participating students from NTNU, where 20 of the replies included an optional textual feedback. 27 (67.5%) of the participants had not been on an exchange program, while 13 (32.5%) had. Figure 6.1 shows the faculty demographic of the participants.

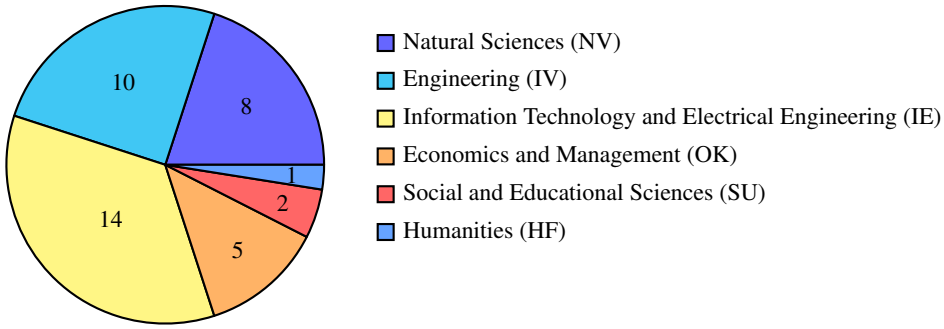


Figure 6.1: Faculty demographic of participants in Questionnaire 2

The results from the closed questions are presented first, followed by the findings from the open question which is summarized in short. The results from the closed questions are divided into two parts; Part one for results that contribute to answering RQ1 (motivational effect), and part two for those that contribute to RQ2 (suitability of CBR).

6.2.1 Part 1: Motivational Effect and Use

The following figures display results from the questions concerning the motivational effect Utsida may have on students to apply for an exchange program. The degree of agreement ranges from *1: Strongly disagree* to *5: Strongly agree* as a Likert-type question.

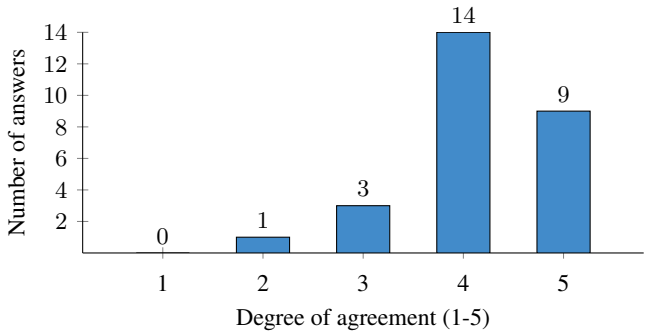


Figure 6.2: Statement: ”I think my motivation for exchange would increase if Utsida was in use”.
Group: Students who have **not** participated in an exchange program. (N = 27)

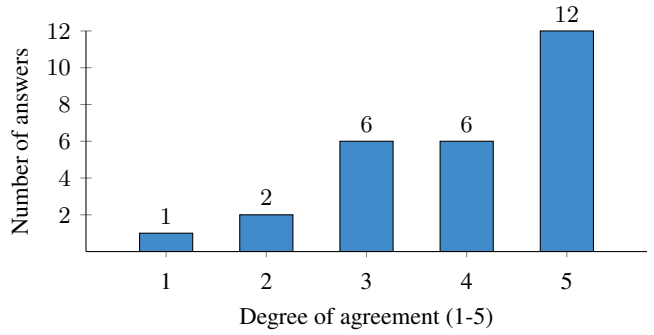


Figure 6.3: Statement: ”I thought it was easy to use Utsida”.
Group: Students who have **not** participated in an exchange program (N = 27)

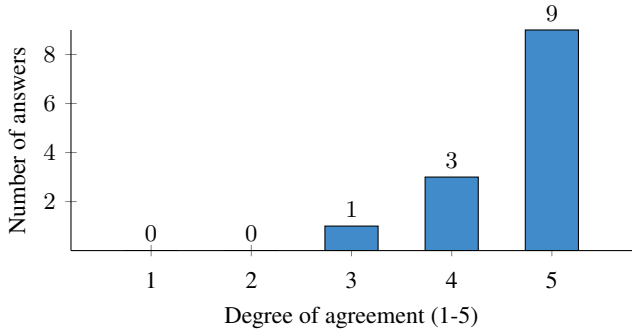
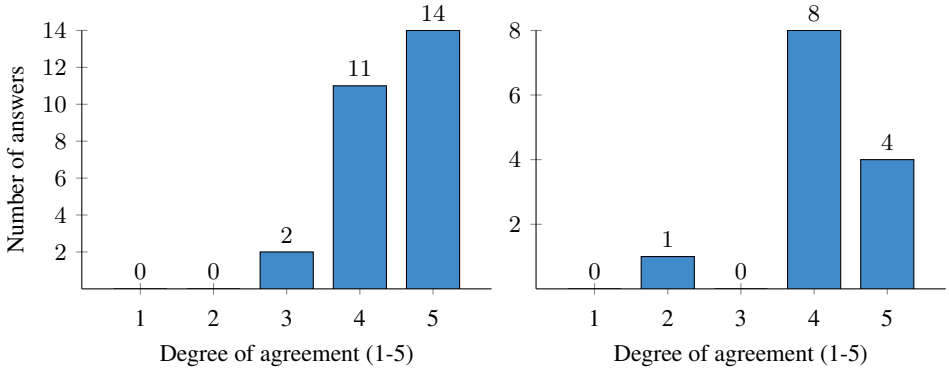


Figure 6.4: Statement: ”I think Utsida would have simplified my application process”.
Group: Students who have participated in an exchange program (N = 13).

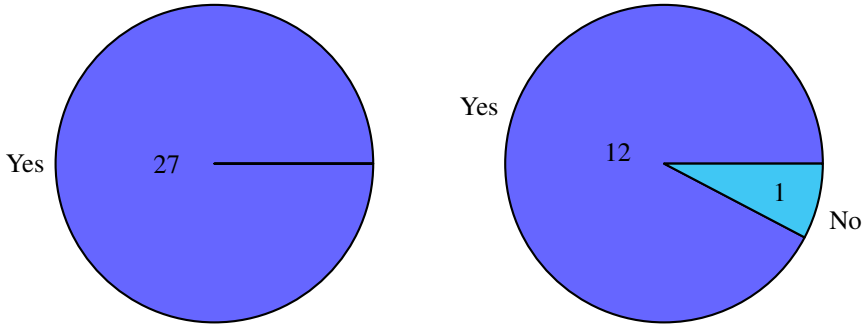


(a) Group: Students who have **not** participated in an exchange program (N = 13)

(b) Group: Students who **have** participated in an exchange program (N = 27)

Figure 6.5: Statement: ”I think Utsida can contribute to an increased number of students who choose to participate in an exchange program”

To make an inference on whether OIR should include Utsida in their exchange program application process, one question asked if the participants would recommend Utsida to the OIR. The results are shown in Figure 6.6 and 39 out of 40 participants would have recommended that OIR included Utsida in the process.



(a) Group: Students who have **not** participated in an exchange program (N = 27)

(b) Group: Students who **have** participated in an exchange program (N = 13)

Figure 6.6: Results from question: "Would you recommend The Office of International Relations to include this system in the exchange application process?"

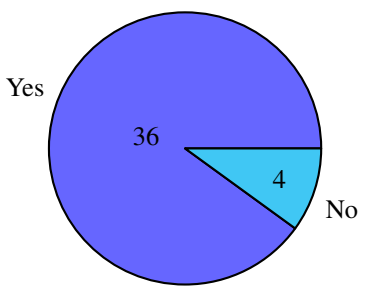
The summary in Table 6.2 shows the mode and frequencies for all the Likert-type questions. Each frequency also has a percentage value indicating the share of answers on the degree of agreement. In addition, the two target groups of questions has a Cronbach's Alpha value calculated to evaluate their internal consistency.

Table 6.2: Summary of results from Likert-type questions

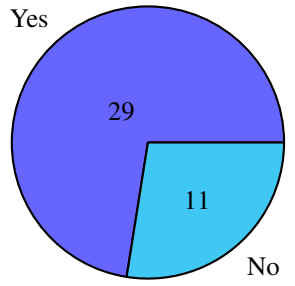
Statement	Mode	Frequencies (degree of agreement) 1	2	3	4	5
Students that **have not** been on exchange		(Cronbach's Alpha = 0.916)				
I think my motivation for exchange would increase if Utsida was in use	4	0	1 (3.7%)	3 (11.1%)	14 (51.9%)	9 (33.3%)
I thought it was easy to use Utsida	5	1 (3.7%)	2 (7.4%)	6 (22.2%)	6 (22.2%)	12 (44.4%)
I think Utsida can contribute to an increased number of students who choose to participate in an exchange program	5	0	0	2 (7.4%)	11 (40.7%)	14 (51.9%)
Students that **have been** on exchange		(Cronbach's Alpha = 0.855)				
I think Utsida would have simplified my application process	5	0	0	1 (7.7%)	3 (23.1%)	9 (69.2%)
I think Utsida can contribute to an increased number of students who choose to participate in an exchange program	4	0	1 (7.7%)	0	8 (61.5%)	4 (30.8%)

6.2.2 Part 2: Evaluation of Recommendations

The following figures display the results from questions that concerned the evaluation of recommendations received in Utsida. The first two questions asked participants to evaluate the recommendations they received, where the query was decided by the user's themselves.



(a) Question: Did the system recommend one or more **universities** which are relevant for you?

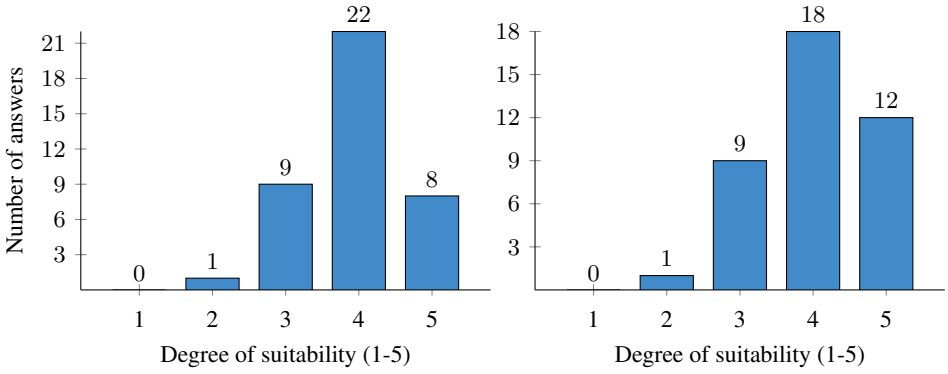


(b) Question: Did the system recommend one or more **courses** which are relevant for you?

Figure 6.7: Results from question asking to evaluate the recommendation on participants' self made query. (N = 40).

Two predefined queries, *query 1* and *query 2*, were evaluated by the participating students. The queries can be seen in Appendix B.1. The degree of suitability ranges from *1: Very*

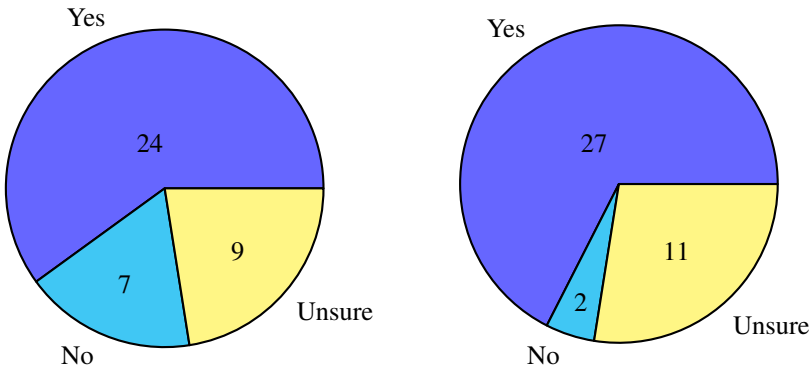
poor to *5: Very good*.



(a) Results for **Query 1**

(b) Results for **Query 2**

Figure 6.8: Results from question: "How did the recommendation suit the original search parameters?" (N = 40)



(a) Question: Did the system recommend relevant courses for one who studies **computer technology?**

(b) Question: Did the system recommend relevant courses for one who studies **electronics?**

Figure 6.9: Results from question on relevancy of recommendation from **(a)** Query 1 and **(b)** Query 2 (N = 40)

6.2.3 Open Question

The open question in questionnaire 2 asked if the participant had any other comments on the site or process in general. 20 participants answered the question. A qualitative theme

analysis [26] was performed and the findings concluded to be mostly positive with 11 out 20 having a positive theme. Five of the participant comments mentioned the site was non-intuitive, four thought it was easy to use, and three wrote Utsida could be time-saving. The full results of the theme analysis can be viewed in Appendix B.2.

6.3 Offline Experiment on the CBR-RS

By using the score matrix given by Table 4.1, 20 different simulated user made queries (App. C.1) were evaluated. Each query was scored (App. C.1) on two different concept configurations, one being the similarity based retrieval used in the CBR-RS and the other being a simulated exact match search. The central tendency results are shown in Table 6.3.

Table 6.3: Result from offline experiment, N=20

	Mean score	Std. Deviation	Std. Error Mean
Similarity based retrieval	36.70	4.38	.978
Simulated exact match search	30.05	3.93	.878

Table 6.4 shows the results from the Paired T-test performed on the two result data sets from the offline experiment. The Sig. (2-tailed) value is the p-value. This value is lower than 0.05, indicating a significant difference between the two sample means.

Table 6.4: Result of the Paired T-test

			95 % Confidence Interval of the Difference				
Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
6.65	3.00	.671	5.244	8.056	9.897	19	0.000

Chapter 7

Discussion

This chapter interprets the results presented in the previous chapter and discusses them with regards to the research questions and goals. Identified limitations of the project are also presented.

7.1 Motivational Effect

Goal 1 was to "Create a prototype that improves the motivation for students at NTNU to apply for an exchange program". This study created that prototype and named it Utsida. Utsida's motivational effect was measured by conducting a user study with an accompanying questionnaire. The questionnaire shows a strong positive influence on motivation, with 23 out of 27 participating students that had not been on exchange agreeing (i.e. answers of either 4: Agree or 5: Strongly agree) that their motivation would likely increase if Utsida was in use. The frequencies also shows that the answers did not vary from the mode of 4, indicating that most of the participants had similar views. More than half of the participants (18 out of 27) further reviewed and agreed that Utsida was easy to use. Utsida should ideally have a higher agreement on the usability, but it is considered acceptable because of the early prototype stage of Utsida. Furthermore, a Cronbach's alpha value of 0.916 on

the Likert scale of participating students that have not been exchange, and 0.855 for those that had, shows a strong sign of internal consistency in questionnaire 2, with values over 0.7 considered satisfactory [6].

The mode for all questions in the first part of questionnaire 2 resides between *4: Agree* and *5: Strongly Agree*. These statistics implies that the participants had a largely positive experience with Utsida. Furthermore, 37 out of the 40 participating students agreed that Utsida could contribute to an increased numbers of students that participate in an exchange program. These results correlate with the results of a study done by NTNU's Office of International Relations (OIR) [27] where exchange students were asked what factors could contribute to an increased number of exchange students. Utsida specifically targets the top factors in the study with 85% of the participants selecting "lists of previously approved courses" and 80% selecting "lists over recommended universities with approved courses". 39 out of 40 participants of questionnaire 2 would also recommend OIR to include Utsida in their application process. This indicates a high user confidence in the system. However, this question could have been especially vulnerable to bias, as further discussed in the limitations (sec. 7.3).

The results from questionnaire 1 concluded that the most important factors of motivation is the study language, social quality and academic quality. These results are supported by the the results of the *push-pull* factors of Mazzarol and Soutar [22] which identifies academic quality, social links and language as important influences. The six most important factors identified in questionnaire 1 were included in the CBR-RS to ensure high relevance of attributes. The open question also identified that general information and an easier application process are the important areas to focus on. These results further increase the confidence in that Utsida, which targets these areas, would have a positive effect on students' motivation for applying for an exchange program.

7.2 Suitability of CBR Methodology

Goal 2 was to "Use the CBR methodology to give relevant recommendations on exchange program universities and courses". The goal was accomplished by creating a CBR-RS that uses myCBR and modeling the exchange experience concept by: converting exchange students' experience reports to cases, setting appropriate similarity measures based on literature and user testing, and giving the attributes weights based on the results of questionnaire 1. The offline experiment concluded that using similarity-based recommendation give more relevant results than traditional exact matching searching. The conducted paired T-test on the results had a p-value less than 0.05, indicating that the results are statistically significant and that the difference is highly unlikely to be created by chance. The closest thing NTNU currently have to Utsida is OIR's site[1] for publishing and finding students' experience reports from exchange programs. This site uses traditional exact matching techniques to search through the reports. Thus, based on the results of the offline experiment, it should be easier to find relevant experiences in Utsida than in OIR's search system.

The results of the '*evaluation of recommendations*' part of questionnaire 2 are highly positive. The results shows that the majority of the participants received relevant recommendations with their queries. However, a clear tendency is that more participating students received relevant recommendations on universities (36 out of 40) than courses (29 out of 40). This tendency might be due to universities being a far more general factor than courses. A university can be of interest for many, while a course is more strictly connected to a particular field of study and year. Furthermore, the courses in a recommendation are considered novel recommendations, which are recommendations for items the user did not know about [32]. Therefore, it could be difficult for the participant to evaluate the relevancy of unknown courses. Since the experience reports also are mostly written in free text, some course titles may be in unknown foreign languages. In addition, the courses do not have any other description than the title, which makes it more difficult for students to know if the courses are relevant for them.

[1] Experience report database: `https://www.ntnu.no/studier/studier_i_utlandet/rapport/`

The results of the questions asking the participants to evaluate the recommendations for the two predefined queries show highly positive tendencies. The relevancy of the recommendations, as a whole, were deemed suitable (i.e. answered either 4: Good or 5: Very good) by 30 out of 40 participants for both queries. This shows that the CBR-RS, in most cases, gives relevant recommendations for a input query. Furthermore, among the 40 participants, 24 thought Utsida recommended relevant courses for the first predefined query, and 27 for the second predefined query. These results suggests that the CBR-RS would likely recommend relevant courses to students similar to the predefined queries. An average of 10 out 40 participants were unsure to deem the recommended courses relevant for each query. This uncertainty could come from many participants not having enough knowledge on the field of study, and are, therefore, not able to evaluate the relevancy of the courses. In hindsight, the questions could have been constructed without the "Unsure" option, which do not contain much value for this research project. However, removing the option might have introduced more bias in the "Yes/No" answers as the participant might be unsure what to select.

7.3 Limitations

Some limitations with the conducted research was discovered during the project duration. The most significant limitation was the small sample size for the questionnaires. Even though both questionnaires got a reasonable amount of response, questionnaire 1 with 84 answers and questionnaire 2 with 40 answers, they were both below the goal and optimal sample size. A confidence level of 95% resulted in a high margin of error of 10.6% on questionnaire 1 and a even higher margin of error of 15.5% on questionnaire 2. However, the result data will show a positive tendency even in the worst case of the population's answers being on the low end of the margin.

Utsida had 165 unique registered users during the user study, but only 40 of them (24%) completed the full test and answered questionnaire 2. This suggests that the user study may have been too complex and time-consuming and therefore made many participants withdraw before finishing. Furthermore, the questionnaires should have undergone more

rigorous pilot testing with experts on the field of questionnaire design to further reduce the risk of bias influence. The question on whether the participant would recommend OIR to include Utsida in their exchange application process was found to be especially vulnerable to acquiescence bias and should have been designed differently.

7.4 Research Questions

RQ1: "*What effect would an information system for recommending and assisting exchange course selection have on students' motivation for doing a study abroad- or student exchange program at NTNU?*"

The results from evaluating Utsida, which is an information system that recommends and assists the exchange course selections for students, show a highly positive effect on students' motivations for doing an exchange program. In addition, the results also indicate that Utsida may, if used, lead to more students applying for exchange programs. The majority of the answers in questionnaire 2 were positive, with an average agreement on statements regarding motivation of 85.8%. Most noteworthy, 37 out of 40 participating students agreed that Utsida could increase the number of students that apply for an exchange program and 23 out of 27 participants that had not been on exchange agreed that their motivation for exchange would increase with Utsida in use.

RQ2: "*How suitable is the case-based reasoning methodology for recommending relevant universities and courses for a study abroad- or student exchange program?*"

The positive responses in the user study, and the significant results of the offline experiment, indicates that the CBR methodology is suitable for giving recommendations on courses and universities. The majority of participating students in the user study received relevant recommendations in Utsida. In detail, 36 out of 40 participants reviewed the university recommendations to be relevant, while 29 out of 40 participants received relevant recommendations on courses. However, the recommendations tended to be used more as a

tool for course inspiration than to give the definitive solutions. This is also more in accordance with giving recommendations; a recommendation may not have a suitable solution, but the user is left with new inspiration and ideas.

Chapter 8

Conclusion

This chapter concludes the project and presents possible future work on the subject.

8.1 Conclusion

This research project has found that a system for recommending courses and universities would positively influence the motivation of students for going on an exchange program and could consequently lead to more students participating in the programs. The influence claim is based on the highly positive user feedback received in the evaluation of Utsida, with an average agreement on statements regarding the motivation of 85.8% (confidence level: 95%, 0margin of error: 15.5%). While this margin of error is relatively high due to the less than optimal sample size, the results of this project still indicate a beneficial effect.

Furthermore, CBR is found to be suitable for giving recommendations on universities and courses; the CBR-RS gave relevant recommendations for most (36 out of 40) of the user study participants, and performed significantly better than standard exact match search. While the recommendations alone can not fully be used as an independent solution, they might give students enough inspiration and motivation to initiate the process. If applied in

a real setting Utsida would also, with its simplification of the course application process, likely reduce the time used for both students and advisers when approving courses for exchange programs. The simplified and faster application process would lead to more available time for advisers to help students in actual need of help and less time used on students that only need to complete the approval process.

Based on the results that show a high approval of Utsida, it is highly recommended that NTNU invests further in digitizing the whole- or part of the application process for exchange programs. An investment could help NTNU reach their goal of 40% of the students participating in exchange programs and contribute to the overall goal in Norway of 20% student mobility. This project has also shown that exchange reports submitted by students can be used in other useful ways. Therefore the data from the experience reports should be public to use (e.g. as an API), and formatted appropriately so that students and others can utilize them in ways that may give a benefit to NTNU. The contributions of this project are not only beneficial to NTNU alone, but could also be used as an inspiration for similar projects at other educational institutions.

8.2 Future Work

Even though the Utsida prototype was developed to a degree needed to answer this research project, it could be extended with new features, and future research could focus on different approaches of evaluation. This section proposes new features in Utsida, possible improvements to the CBR-RS, and ideas for further research.

8.2.1 Web Application Features

Several suggestions for improving or expanding features of Utsida emerged throughout the development process by reviewing: interviews with exchange course advisers, feedback from usability testing, and textual answers from the questionnaires. Many of the features were outside the scope of this research project and were not implemented; the following list describe the most noteworthy ones:

Let advisers make comments on course approval applications: Currently, a course application in Utsida is only either approved or disapproved with no comments from advisers. The advisers often wants to include special premises for courses to be approved. By implementing a comment functionality for advisers, advisers could give premises as comments on applications, and through these comments clarify what the student should do to to get the courses approved.

Expiration on approval of course matches: Course matches tend to be irrelevant if they are too old, due to changes in the content of the courses. It could therefore be useful to automatically flag old approvals.

Implement filters on the course match list: The list of course matches currently only have search functionality. The list functionality could be expanded to include filtering on institute or faculty, making it easier to find relevant course matches.

Support for multiple departments: Some study programs focus on many fields of study, and therefore belongs to multiple departments. Support for multiple departments could be implemented in Utsida by letting users register with multiple departments. The CBR-RS could for the recommendations give partial similarity in experiences with either of the user's departments.

Interactive map of exchange occurrences: The map on the home page of the web application only display the number of students who have been on exchange in a specific country. This map could be extended to be more interactive, for example by adding support of selecting a specific country to get recommendations.

8.2.2 Possible Extensions to the CBR-RS

The CBR-RS could be further improved by adding more relevant attributes to the concept and by including adaption and learning techniques to personalize and further increase the relevancy of recommendations.

Expanding and Improving the Exchange Experience Concept Several attributes that were mentioned in the feedback from the usability tests and questionnaires were not included in the exchange experience concept. These suggestions included: cost of living, weather and climate, tuition fees, university rankings, and whether a trip was arranged through a specific program such as Erasmus[1]. None of these are possible to extract from the experience reports, so the data would have to come from external APIs. However, if good sources for such data is found, adding more attributes to the concept would be beneficial for the CBR-RS to recommend more personalized and relevant exchange experiences. For example, Numbeo[2] offers a universal API for cost of living in all countries and cities, and could be used to add cost of living for the country in each experience.

Adaption in the CBR-RS Utsida currently has no form of learning or adaption. Utsida could be extended to learn the preferences of a specific user by requiring feedback on recommendations or by asking for more user information. It could then use this data to increase the relevancy of the recommendations given to the user and adapt the results specifically to the user. For example, if a specific recommendation is a good match for a user, but it only includes a few courses, it could merge its courses with the courses in other similar recommendations. The recommendation would then become more complete.

Automatic Parsing of New Exchange Experiences To automatically expand the case-base in Utsida, it would have to parse and add new experience reports as they are published on OIR's site. This could be beneficial to ensure Utsida always has the latest and most relevant data. However, to enable this functionality, Utsida would have to be configured to accept the format of the new experience reports system released in 2017.

[1]Erasmus: a European Union student exchange program
[2]Numbeo: `https://www.numbeo.com/cost-of-living/`

8.2.3 Further Research

If Utsida is further developed to work seamlessly in a production environment, it can be possible to evaluate the system in a real-life situation. The following suggestions could be possible ways to study how Utsida affects motivation in real-life.

Testing Utsida in an Application Process Utsida could be further evaluated by testing it in a real application process for a collection of students. Data could then be collected to find out what these students thought of using Utsida in the application process and compare it to the students who went through the standard process. The resulting data could be used to further argue that Utsida should be included in the OIR's process, and further document its effect on motivation.

Observing Real Effects of Utsida Utsida could be further evaluated by including the system fully in the application process for an exchange program at NTNU. It would then be of value to compare the number of students applying before and after the inclusion of Utsida. If the number of students applying for an exchange program increased significantly it could be argued that a system like Utsida has a definitive effect on the number of students that go on an exchange program.

Bibliography

[1] A. Aamodt and E. Plaza, "Case-based reasoning: Foundational issues, methodological variations, and system approaches," *AI communications*, vol. 7, no. 1, pp. 39–59, 1994.

[2] I. E. Allen and C. A. Seaman, "Likert scales and data analyses," *Quality progress*, vol. 40, no. 7, p. 64, 2007.

[3] N. Askham, D. Cook, M. Doyle, H. Fereday, M. Gibson, U. Landbeck, R. Lee, C. Maynard, G. Palmer, and J. Schwarzenbach, "The six primary dimensions for data quality assessment," Technical report, DAMA UK Working Group, Tech. Rep., 2013.

[4] S. Begum, M. U. Ahmed, P. Funk, N. Xiong, and M. Folke, "Case-based reasoning systems in the health sciences: a survey of recent trends and developments," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 41, no. 4, pp. 421–434, 2011.

[5] R. Bergmann, *Experience management: foundations, development methodology, and internet-based applications*. Springer-Verlag, 2002.

[6] J. M. Bland and D. G. Altman, "Statistics notes: Cronbach's alpha," *The BMJ*, vol. 314, no. 7080, p. 572, 1997.

[7] B. W. Boehm, “A spiral model of software development and enhancement,” *Computer*, vol. 21, no. 5, pp. 61–72, 1988.

[8] F. Bousbahi and H. Chorfi, “Mooc-rec: a case based recommender system for moocs,” *Procedia-Social and Behavioral Sciences*, vol. 195, pp. 1813–1822, 2015.

[9] R. Bramucci and J. Gaston, “Sherpa: increasing student success with a recommendation engine,” in *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. ACM, 2012, pp. 82–83.

[10] U. Brandenburg, S. Berghoff, and O. T. Álvarez, “The erasmus impact study: Effects of mobility on the skills and employability of students and the internationalisation of higher education institutions,” *Publications Office of the European Union, Luxembourg*, 2014.

[11] D. Bridge, M. H. Göker, L. McGinty, and B. Smyth, “Case-based recommender systems,” *The Knowledge Engineering Review*, vol. 20, no. 03, pp. 315–320, 2005.

[12] J. Brooke *et al.*, “Sus-a quick and dirty usability scale,” *Usability evaluation in industry*, vol. 189, no. 194, pp. 4–7, 1996.

[13] L. J. Cronbach, “Response sets and test validity,” *Educational and psychological measurement*, vol. 6, no. 4, pp. 475–494, 1946.

[14] D. R. Fesenmaier, F. Ricci, E. Schaumlechner, K. Wöber, C. Zanella *et al.*, *DIETORECS: Travel advisory for multiple decision styles*. na, 2003.

[15] P. C. Fishburn, “Letter to the editor—additive utilities with incomplete product sets: application to priorities and assignments,” *Operations Research*, vol. 15, no. 3, pp. 537–542, 1967.

[16] A. Furnham, “Response bias, social desirability and dissimulation,” *Personality and individual differences*, vol. 7, no. 3, pp. 385–400, 1986.

[17] M.-J. Huang, M.-Y. Chen, and S.-C. Lee, “Integrating data mining with case-based reasoning for chronic diseases prognosis and diagnosis,” *Expert systems with applications*, vol. 32, no. 3, pp. 856–867, 2007.

[18] R. V. Krejcie and D. W. Morgan, “Determining sample size for research activities,” *Educational and psychological measurement*, vol. 30, no. 3, pp. 607–610, 1970.

[19] A. K.-y. Leung, W. W. Maddux, A. D. Galinsky, and C.-y. Chiu, “Multicultural experience enhances creativity: the when and how.” *American Psychologist*, vol. 63, no. 3, p. 169, 2008.

[20] R. Likert, “A technique for the measurement of attitudes.” *Archives of psychology*, 1932.

[21] J. María Cubillo, J. Sánchez, and J. Cervino, “International students’ decision-making process,” *International Journal of Educational Management*, vol. 20, no. 2, pp. 101–115, 2006.

[22] T. Mazzarol and G. N. Soutar, ““push-pull” factors influencing international student destination choice,” *International Journal of Educational Management*, vol. 16, no. 2, pp. 82–90, 2002.

[23] S. Mulyana, S. Hartati, R. Wardoyo, and E. Winarko, “Case-based reasoning for selecting study program in senior high school,” *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 4, pp. 136–140, 2015.

[24] D. A. Norman and S. W. Draper, “User centered system design,” *New Perspectives on Human-Computer Interaction, L. Erlbaum Associates Inc., Hillsdale, NJ*, vol. 3, 1986.

[25] Norwegian Centre for International Cooperation in Education (SIU), “Studentutveksling fra norge,” 2017. [Online]. Available: https://www.siu.no/publikasjoner/Alle-publikasjoner/Studentutveksling-fra-Norge.-SIU-rapport-02-2016

[26] B. J. Oates, *Researching information systems and computing*. Sage, 2005.

[27] Office of international Relations NTNU, “Survey on information exchange students receive from NTNU,” 2016, Unpublished internal survey.

[28] C. E. Osgood, “The nature and measurement of meaning.” *Psychological bulletin*, vol. 49, no. 3, p. 197, 1952.

[29] L. Quijano-Sánchez, J. A. Recio-García, B. Díaz-Agudo, and G. Jiménez-Díaz, "Happy movie: A group recommender application in facebook," in *24th International Florida Artificial Intelligence Research Society Conference, FLAIRS*, 2011, pp. 127–134.

[30] M. M. Richter and R. Weber, *Case-based reasoning: a textbook*. Springer Science & Business Media, 2013.

[31] T. Roushan, D. Chaki, O. Hasdak, M. S. Chowdhury, A. A. Rasel, M. A. Rahman, and H. Arif, "University course advising: Overcoming the challenges using decision support system," in *Computer and Information Technology (ICCIT), 2013 16th International Conference on*. IEEE, 2014, pp. 13–18.

[32] G. Shani and A. Gunawardana, "Evaluating recommendation systems," in *Recommender systems handbook*. Springer, 2011, pp. 257–297.

[33] A. Stahl and T. R. Roth-Berghofer, "Rapid prototyping of cbr applications with the open source tool mycbr," in *European Conference on Case-Based Reasoning*. Springer, 2008, pp. 615–629.

[34] S. Trewin, "Knowledge-based recommender systems," *Encyclopedia of library and information science*, vol. 69, no. Supplement 32, p. 180, 2000.

[35] E. Triantaphyllou, "Multi-criteria decision making methods," in *Multi-criteria Decision Making Methods: A Comparative Study*. Springer, 2000, pp. 5–21.

[36] V. Vaishnavi and W. Kuechler, *Design research in information systems*. Springer, 2004.

[37] C. Vialardi, J. Bravo Agapito, L. S. Shafti, and A. Ortigosa, *Recommendation in higher education using data mining techniques*. Barnes, T., Desmarais, M., Romero, C., & Ventura, S., 2009.

[38] T. R. Williams, "Exploring the impact of study abroad on students' intercultural communication skills: Adaptability and sensitivity," *Journal of studies in international education*, vol. 9, no. 4, pp. 356–371, 2005.

[39] T. Yamane, “Statistics: an introductory analysis-3,” 1973.

[40] W.-S. Yang and S.-Y. Hwang, “iTravel: A recommender system in mobile peer-to-peer environment,” *Journal of Systems and Software*, vol. 86, no. 1, pp. 12–20, 2013.

[41] N. J. Yuan, Y. Zheng, L. Zhang, and X. Xie, “T-finder: A recommender system for finding passengers and vacant taxis,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 10, pp. 2390–2403, 2013.

Appendix

Appendix A

Questionnaire 1: Motivational Factors

A.1 Results

Når du reiste/eventuelt reiser på utveksling, hvor avgjørende er/var følgende faktorer for deg?

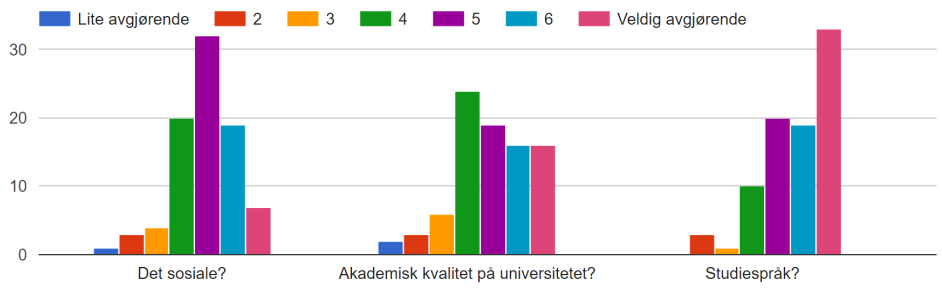


Figure A.1: Motivational factors, part 1

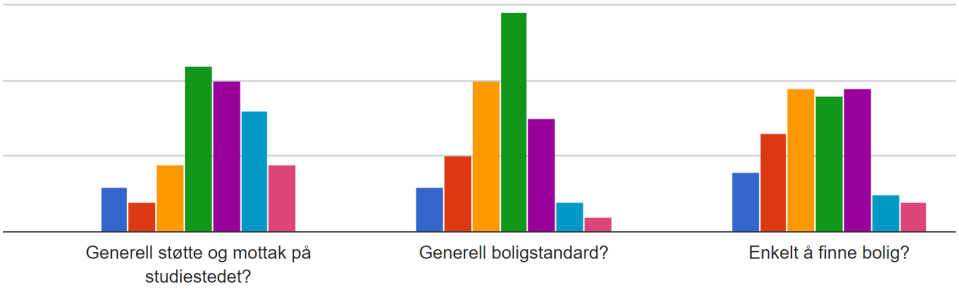


Figure A.2: Motivational factors, part 2

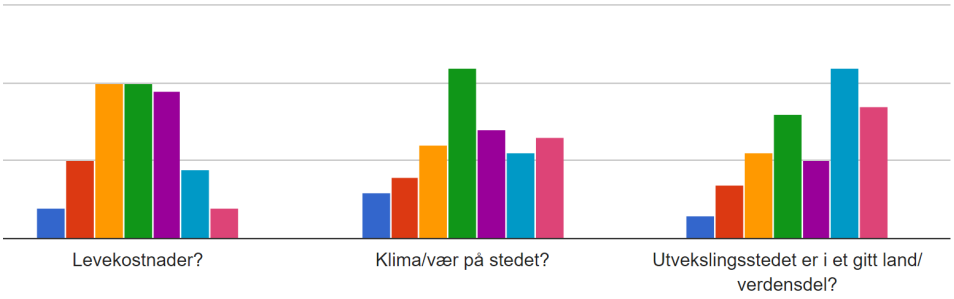


Figure A.3: Motivational factors, part 3

Cate	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20	#21	#22	#23	#24	#25	#26	#27	#28	#29	#30	#31	#32	#33	#34	#36	SUM
Fag	*						*	*		*	*	*	*							*		*	*					*				*				12
Informasjon								*	*	*		*			*																	*				6
Søknadsp	*	*	*			*				*										*																6
Kultur					*	*												*	*											*						5
Opplevelser						*															*			*	*											4
Sosial integrasjon																			*		*				*					*						4
Språk																*	*											*								3
Geografisk lokasjon	*																*													*						3
Veiledning												*			*																					2
Bymiljø																										*		*								2
Skolepenger																											*		*							2
Gruppe																											*								*	2
Ranking																																*		*		2
Semesterform														*																						1
Jobb																	*																			1
Erasmus																															*					1
Sentral lokasjon																																*				1
Norske studenter																																*				1

Figure A.4: List of word frequencies from open question

A.2 Questions

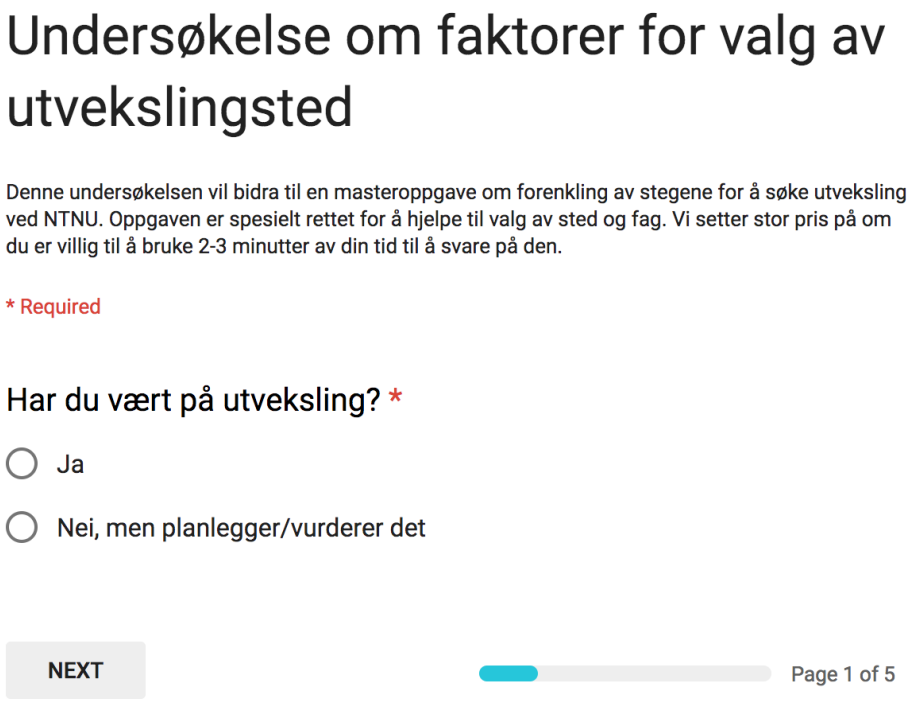


Figure A.5: Introduction to questionnaire 1

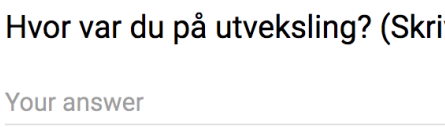


Figure A.6: Question 2: Where did you go on exchange?

Figure A.7: Question 3: What university do you belong to?

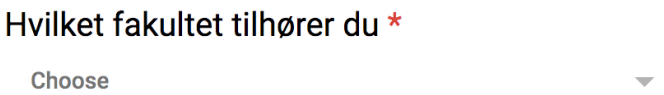


Figure A.8: Question 4: If NTNU, what faculty do you belong to?

Når du reiste/eventuelt reiser på utveksling, hvor avgjørende er/var følgende faktorer for deg? *

Ranger hvor avgjørende hver faktor er/var for din beslutning av studiested

	Lite avgjørende	2	3	4	5	6	Veldig avgjørende
Det sosiale?	○	○	○	○	○	○	○
Akademisk kvalitet på universitetet?	○	○	○	○	○	○	○
Studiespråk?	○	○	○	○	○	○	○
Generell støtte og mottak på studiestedet?	○	○	○	○	○	○	○
Generell boligstandard?	○	○	○	○	○	○	○
Enkelt å finne bolig?	○	○	○	○	○	○	○
Levekostnader?	○	○	○	○	○	○	○
Klima/vær på stedet?	○	○	○	○	○	○	○
Utvekslingsstedet er i et gitt land/verdensdel?	○	○	○	○	○	○	○

Er det noen andre faktorer som påvirket valget ditt, eller du synes er viktig?

Your answer

Figure A.9: Question 5: Main question, rate the factors. And Question 6: The open question.

Appendix B

Questionnaire 2: User Study of Utsida

B.1 Predefined Queries

Table B.1: Predefined search queries used in questionnaire 2

Query	Query 1	Query 2
Institute	Institutt for elkraftteknikk	Institutt for datateknikk og informasjonsvitenskap
Continent	North America	Europe
Country	USA	Frankrike
University		
Language	Engelsk	Fransk
Study Period	2017	2017
Academic Quality	6	8
Social Quality	9	4
Residential Quality	7	6
Reception Quality	3	3

B.2 Open Question Results

Table B.2: Theme Analysis

Category	#1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	SUM
Positive feedback	*	*				*		*	*				*	*			*	*	*	*	11
non-intuitive				*			*			*				*	*						5
Easy to use		*			*							*			*						4
Time saving	*																		*	*	3
Great use of experience reports					*										*						2
Good recommendations	*																				1
Does not fit well for mixed studies							*														1
Great to be self-serviced																			*		1
Idividual adaption																				*	1

B.3 Questions

Undersøkelse etter test av Utsida, en nettside for enklere utveksling

Denne undersøkelsen vil bidra til datagrunnlag for en masteroppgave som ønsker å gjøre søknadsprosedyren for utveksling ved NTNU lettere. Oppgaven er spesielt rettet mot å forenkle valg av studiested og fag ved å utvikle en prototype vi har navngitt Utsida. Dersom du er student på NTNU setter vi stor pris på om du er villig til å bruke 10-15 minutter av din tid til å teste Utsida og svare på denne tilhørende undersøkelsen.

Undersøkelsen baserer seg på at du har vært innom nettsiden og testet ut de forskjellige funksjonalitetene som finnes. Hvis ikke, så kan du gå til utsida.idi.ntnu.no og prøve den ut. Se punktene under for hva vi ønsker du skal ha gjennomført.

Eventuelle spørsmål om undersøkelsen eller siden kan sendes til utsida.ntnu@gmail.com. Nettsiden er tilpasset til enheter med større skjerm som en bærbar pc eller stasjonær.

*Må fylles ut

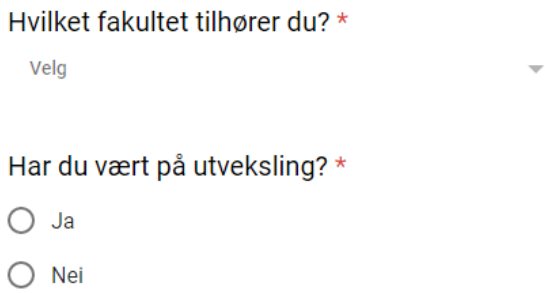
Vennligst bekreft at du har gjennomført følgende punkter på utsida.idi.ntnu.no *

- [ ] Få Anbefaling: Fått anbefaling til universiteter og fag
- [ ] Godkjente fag: Sett på allerede godkjente fag ved ulike universiteter
- [ ] Mine fag: Koblet NTNU-fag og fag fra utlandet til en fagkobling
- [ ] Mine fag: Opprettet en søknad for godkjenning av fag

NESTE

Side 1 av 7

Figure B.1: The introduction to the questionnaire



Hvilket fakultet tilhører du? *

Velg

Har du vært på utveksling? *

- ◯ Ja
- ◯ Nei

Figure B.2: Question 1: Demographics

Angi i hvor stor grad du er enig eller uenig med påstandene

Jeg tror min motivasjon for utveksling ville økt dersom Utsida var i bruk *

	1	2	3	4	5	
Sterkt uening	○	○	○	○	○	Sterkt enig

Jeg syntes det var enkelt å bruke Utsida *

	1	2	3	4	5	
Sterkt uenig	○	○	○	○	○	Sterkt enig

Jeg tror Utsida kan bidra til at flere velger å dra på utveksling *

	1	2	3	4	5	
Sterkt uenig	○	○	○	○	○	Sterkt enig

Andre kommentarer eller tilbakemeldinger om siden?

Svaret ditt

Ville du anbefalt internasjonal seksjon å inkludere dette systemet i en søknadsprosess for utveksling? *

- ○ Ja
- ○ Nei

Figure B.3: Questions regarding motivational effect on students who have not been on exchange

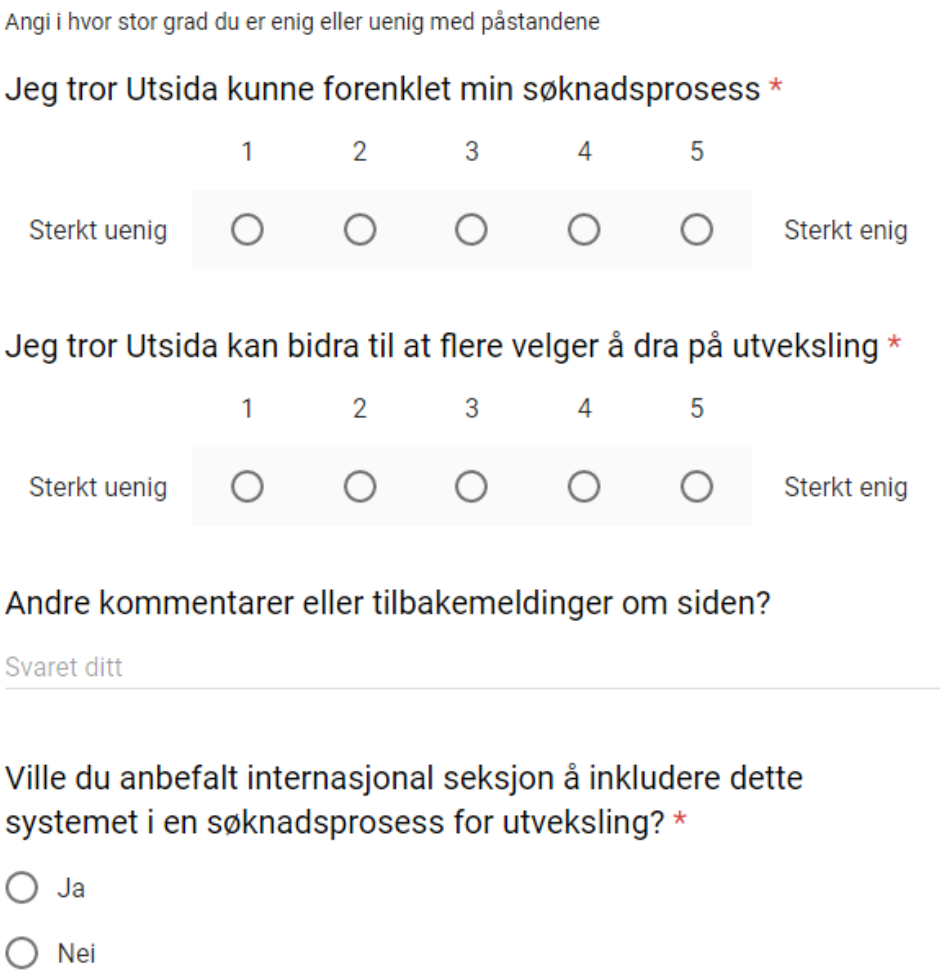


Figure B.4: Questions regarding motivational effect on students who have already been on exchange

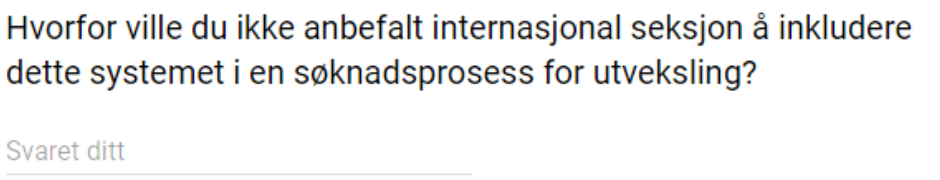


Figure B.5: Question given to participants who answered “No” on the past quetion

Denne delen tar utgangspunkt i at du har gjennomført et søk på utsida.idi.ntnu.no og fått anbefaling på universiteter og fag.

Anbefalte systemet ett eller flere universiteter som var relevante for deg? *

- ◯ Ja
- ◯ Nei

Anbefalte systemet ett eller flere fag som var relevante for deg? *

- ◯ Ja
- ◯ Nei

Figure B.6: Question asking about the recommendations the participants received during the test

Denne oppgaven går ut på å vurdere en anbefaling for et forhåndsinnstilt søk.

Søk 1 - Gå til utsida.idi.ntnu.no/test1, siden vil da gjennomføre et automatisk søk tilsvarende bildet under.

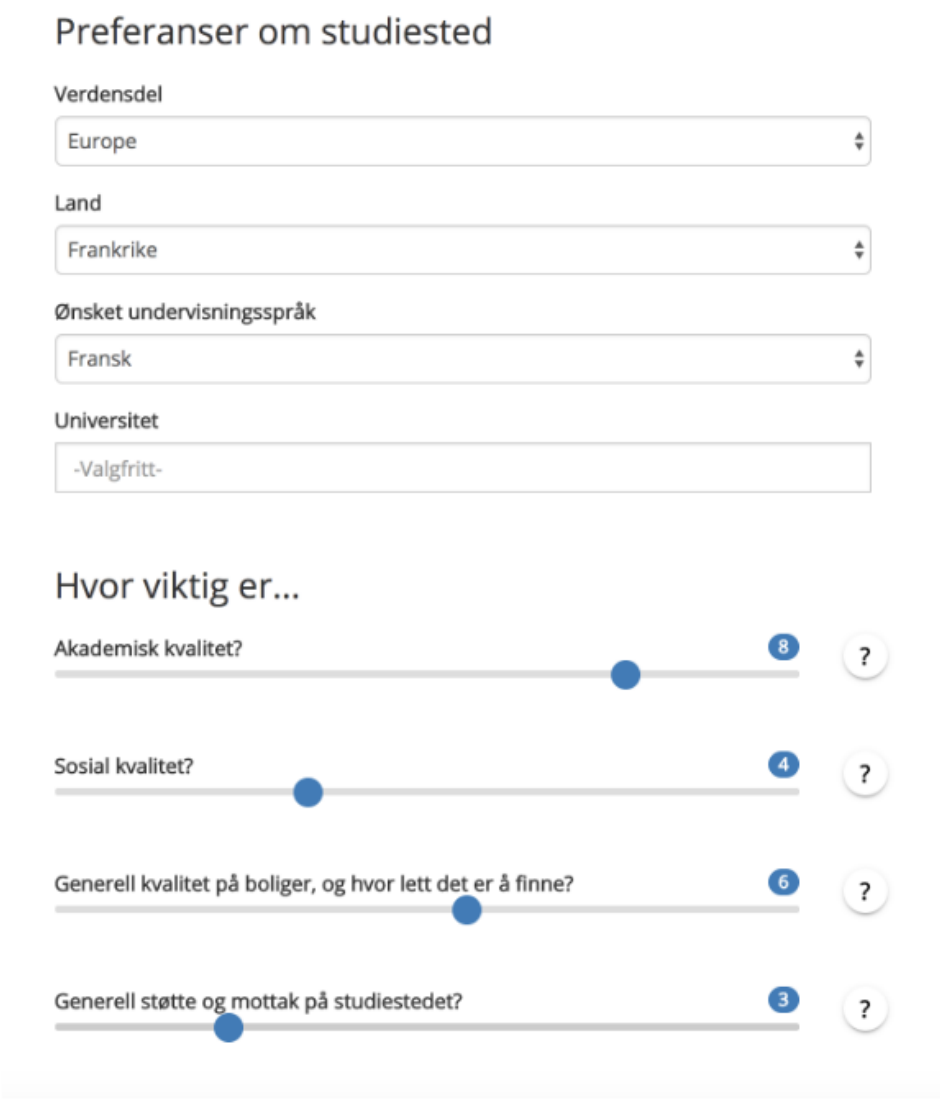


Figure B.7: The first predesigned query

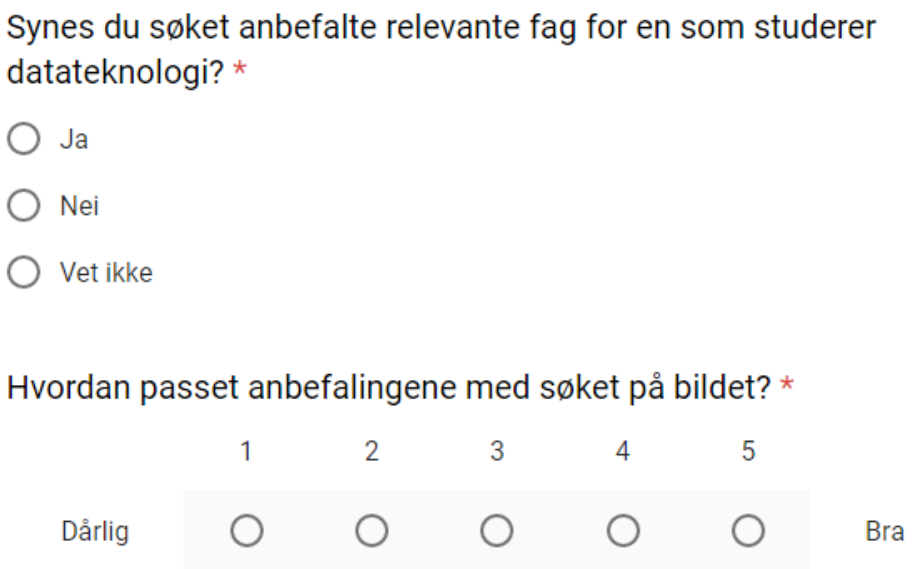


Figure B.8: Evaulation of the first predesigned query

Søk 2 - Gå til utsida.idi.ntnu.no/test2, siden vil da gjennomføre et automatisk søk tilsvarende bildet under.

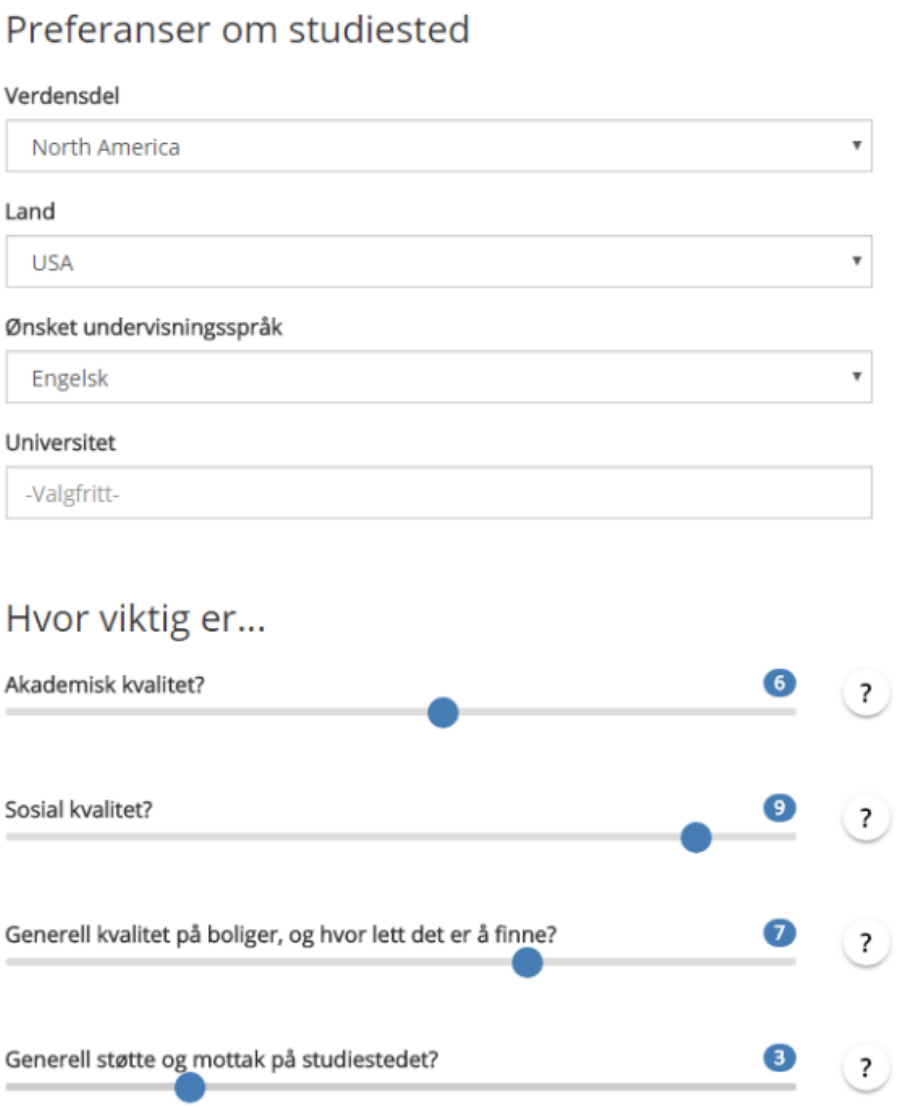


Figure B.9: Second predesigned query

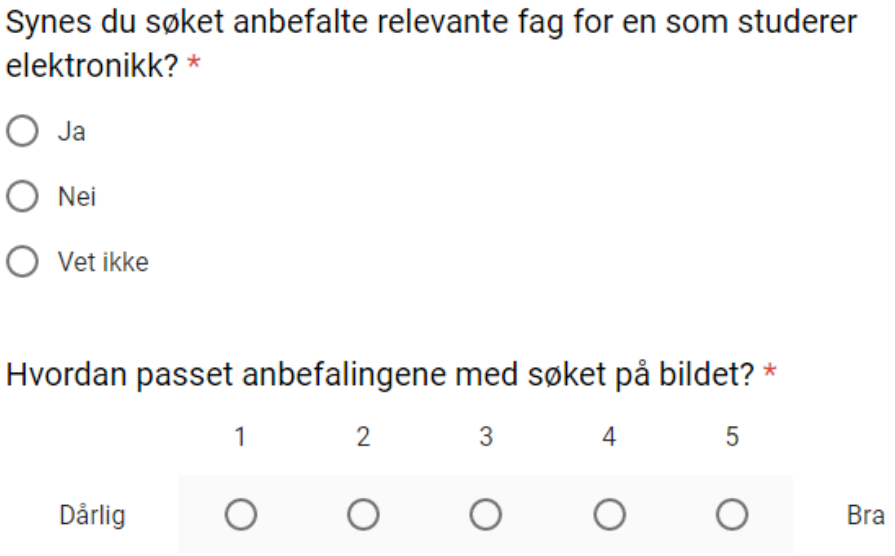


Figure B.10: Evaulation of the second predesigned query

Appendix C

Offline Experiment

C.1 Simulated User Queries

Test-ID	Institutt	Verdensdel	Land	Språk	Universitet	Akademisk	Sosial	Bolig	adm-støtte
1	Institutt for datateknikk og informasjonsvitenskap	North America	USA	Engelsk		6	8	5	4
2	Institutt for datateknikk og informasjonsvitenskap			Engelsk		4	10	4	4
3	Institutt for fysikk	Europe	Storbritannia	Engelsk	University of Cambridge	10	3	5	7
4	Institutt for telematikk	Europe	Tyskland	Tysk		5	5	5	5
5	Institutt for byforming og planlegging	North America	Canada			6	7	7	10
6	Institutt for geologi og bergteknikk	Oceania	Australia	Engelsk		3	8	5	8
7	Institutt for tverrfaglige kulturstudier	Asia	Japan			1	8	8	8
8	Institutt for nevromedisin	Europe	Frankrike	Fransk		10	1	3	10
9	Institutt for pedagogikk og livslang læring	South America	Brazil			8	9	4	8
10	Institutt for byggekunst, historie og teknologi	Oceania	New Zealand	Engelsk		5	5	5	5
11	Psykologisk institutt	Europe	Tyskland	Engelsk	Humboldt Universität	7	6	2	2
12	Institutt for datateknikk og informasjonsvitenskap	Oceania	Australia	Engelsk	Royal Melbourne Institute of Technology	9	8	3	8
13	Institutt for elkrafteknikk	North America	USA			2	10	1	2
14	Institutt for bygg, anlegg og transport	Asia				7	8	3	2
15	Institutt for vann- og miljøteknikk	Europe	Spania	Spansk		4	7	5	10
16	Institutt for datateknikk og informasjonsvitenskap	Asia	Kina	Kinesisk		8	5	8	4
17	Institutt for elektronikk og telekommunikasjon	North America		Engelsk		4	7	1	4
18	Institutt for datateknikk og informasjonsvitenskap	South America	Argentina			4	3	7	9
19	Institutt for sosiologi og statsvitenskap	Africa	Sør-Afrika			6	6	3	3
20	Institutt for byggekunst, prosjektering og forvaltning			Spansk		7	5	9	7

Figure C.1: 20 simulated user queries

C.2 Results

Table C.1: Full results of the offline experiment

Query #	Similarity based retrieval (0-60)	Simulated exact match search (0-60)
1	44.5	32.0
2	34.5	27.0
3	36.5	28.0
4	39.5	32.0
5	41.0	32.5
6	39.5	34.5
7	34.5	28.0
8	41.0	37.5
9	33.5	29.5
10	36.5	26.5
11	35.0	32.5
12	39.0	34.5
13	41.5	32.5
14	36.5	28.5
15	39.0	26.5
16	26.5	20.5
17	36.5	29.5
18	34.5	31.0
19	37.0	33.0
20	27.5	25.0
Mean	36.7	30.05
STD	4.36	3.93

Appendix D

Installation Guide

This guide shows all steps necessary to install and run Utsida locally. The guide is made for Ubuntu 16.04, which is the operative system the Utsida was developed to run on.

Prerequisites

The following sections shows commands which downloads the necessary programs to run this project. Skip the programs you already have installed.

Install Java 8

```
sudo add-apt-repository ppa:webupd8team/java
sudo apt-get update
sudo apt-get install oracle-java8-installer
```

Install Python3 and Pip3

```
sudo apt-get install python3
sudo apt-get install python3-pip
```

Install Git

```
sudo apt-get install git
```

Download the Git Repositories

The following commands downloads the two necessary Git repositories for this project. The first being Utsida's web application, and the second being the MyCBR project.

```
git clone https://github.com/IT3901-Master/utsida.git
git clone https://github.com/IT3901-Master/mycbr-deployment.git
```

Setup and Run

The following commands will install all necessary packages to run the project, initiate all required data, and run it. First, move to the root directory of the utsida project, then run the following commands.

```
sudo pip3 install -r requirements.txt
make initiate
make lrun
```

Open a new terminal or tab, and move to the directory of the mycbr-deployment project. Then run the following command.

```
make run
```