



Aleksandar Bobić, Bsc

Infret - An Interactive Tool for Helping Students Learn and Explore Information Retrieval Concepts

Master's Thesis

to achieve the university degree of
Master of Science

Master's degree programme: Computer Science

submitted to

Graz University of Technology

Supervisor

Gütl, Christian, Assoc.Prof. Dipl.-Ing. Dr.techn.

External Co-supervisors

Dr Christopher Cheong, Dr France Cheong, Dr Justin Filippou

Institute of Interactive Systems and Data Science

Head: Lindstaedt, Stefanie, Univ.-Prof. Dipl.-Inf. Dr.

Graz, December 2019



Aleksandar Bobić, Bsc

Infret - Ein interaktives Tool, zum Vermitteln und Erkunden von Konzepten zu Information Retrieval

Master Arbeit

zur Erlangung des akademischen Grades eines
Diplom-Ingenieur
Masterstudium: Computer Science

eingereicht an der

Technische Universität Graz

Betreuer

Gütl, Christian, Assoc.Prof. Dipl.-Ing. Dr.techn.

Externer Mitbetreuer

Dr Christopher Cheong, Dr France Cheong, Dr Justin Filippou

Institute of Interactive Systems and Data Science

Vorstand: Lindstaedt, Stefanie, Univ.-Prof. Dipl.-Inf. Dr.

Graz, December 2019

Affidavit

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master's thesis.

Date

Signature

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Dissertation identisch.

Datum

Unterschrift

Acknowledgment

First and foremost I would like to thank my mentor Christian Gütl for his guidance, help and understanding during the entire research process. I gained a different outlook on research and many valuable life experiences which would not be possible without you. I would also like to thank my external co-supervisors Christopher Cheong, France Cheong from RMIT, Australia and Justin Filippou from the University of Melbourne who warmly welcomed me to Melbourne, helped me get adjusted and guided me through the research process while staying there and also after coming back to Graz. Additionally, I would like to acknowledge and thank the Graz University of Technology for funding and RMIT University's School of Business IT and Logistics for hosting me as part of my master thesis research.

Thank you to my family for supporting me through my studies, believing in me, motivating me and being there for me through some tough but also many great moments.

I would also like to thank Kristina for supporting me during my studies and for the countless hours spent studying together for various exams. Furthermore, I would like to thank Isabel for being my travel buddy in Australia and for the help she provided while I was working on my thesis. I would also like to thank Adrian for being a positive influence and for motivating me to do more than just study. Additionally, I would like to thank all of my friends that believed in me.

Abstract

An exploratory and interactive prototype called Infret was developed to help students better understand concepts of information retrieval in an information search and retrieval (ISR) class enabled by motivational active learning (MAL). Infret was built using modern web technologies such as Angular and Node.js and based on the findings of a literature survey and the evaluation of a previously built Java-based tool. The Infret prototype enabled students to explore various text statistics from one of the selected text document collections visually and interactively. It was used as part of a class activity by a group of students who were attending the ISR course at TU Graz for a semester. At the end of the activity, the students filled out an evaluation survey. Findings indicate the Infret prototype has above average usability and sparked mostly positive emotions for students. Additionally, multiple improvement possibilities were identified. The aforementioned findings were used to improve and expand Infret. The second iteration of Infret supported various term weighting concepts, a new heatmap visualisation and many other improvements. The second version of Infret was used as part of a class activity by a novice group of students and an experienced group with prior knowledge in IR. The students from both groups filled out an anonymous survey after completing the activities. Findings indicate the second version of Infret helped both groups of students understand various text statistics and term weighting concepts and sparked the novice students curiosity. The usability score given by experienced students was lower than the one received for the Infret prototype, however, it was still higher than the average usability. On the other hand, the usability estimated by novice students was lower than the average. Additionally, even though the negative emotions were more intense than for the Infret prototype the second version of Infret still mainly sparked positive emotions. These and many other findings provide room for further improvements and feature set expansions for future iterations of Infret.

Kurzfassung

Ein explorativer und interaktiver Prototyp namens Infret wurde entwickelt, um Studierenden ein besseres Verständnis über Information Retrieval (IR), das sich mit dem Gewinnen von Informationen beschäftigt, zu ermöglichen. Die Lehrinhalte dazu wurden in einer Information Search and Retrieval (ISR) Lehrveranstaltung mit Hilfe von Motivational Active Learning (MAL) vermittelt. Infret wurde mit modernen Webtechnologien wie Angular und Node.js erstellt, basierend auf den Ergebnissen einer Literaturrecherche und der Evaluierung einer bereits existierenden Java-basierten Softwarelösung. Mit dem Infret Prototyp können Studierende verschiedene Textstatistiken aus einer der ausgewählten Textdokumentensammlungen visuell und interaktiv untersuchen. Infret wurde im Rahmen einer Klassenaktivität von einer Studierendengruppe verwendet, die ein Semester lang am ISR-Kurs teilnahmen. Am Ende der Aktivität füllten die Studierenden eine Bewertungsumfrage aus. Die Ergebnisse zeigen, dass der Infret-Prototyp eine überdurchschnittlich hohe Usability aufweist und bei den Studierenden überwiegend positive Emotionen auslöst. Darüber hinaus wurden mehrere Verbesserungsmöglichkeiten identifiziert. Diese Erkenntnisse wurden zur Verbesserung und Erweiterung von Infret genutzt. Die zweite Version von Infret beinhaltet verschiedene Begriffsgewichtungskonzepte, eine neue Heatmap-Visualisierung und viele andere Verbesserungen. Diese Version wurde im Rahmen einer Klassenaktivität von einer Anfängergruppe und einer Fortgeschrittenengruppe mit Vorerfahrung in IR verwendet. Beide Gruppen haben nach Abschluss der Aktivitäten eine anonyme Umfrage ausgefüllt. Die Ergebnisse deuten darauf hin, dass die zweite Version von Infret beiden Gruppen dabei half, verschiedene Textstatistiken und Begriffsgewichtungskonzepte zu verstehen, und die Neugier der Anfänger weckte. Die von erfahrenen Usern angegebene Usability Wertung war niedriger als jene für die erste Version von Infret, jedoch immer noch höher als die durchschnittliche Usability Wertung. Die von Anfängern bewertete Usability war wiederum niedriger als der Durchschnitt. Obwohl die negativen Emotionen intensiver waren als beim Prototyp, löste die zweite Version noch immer hauptsächlich positive Emotionen aus. Diese und weitere Erkenntnisse bieten Raum für weitere Verbesserungen und Funktionserweiterungen für zukünftige Infret Versionen.

Contents

1. Introduction	1
1.1. Motivation and Focus	2
1.2. Structure	2
2. Background and Related Work	5
2.1. Basic Pedagogic Approaches	5
2.2. Interactive Learning	6
2.2.1. Constructivist theory	7
2.2.2. Experiential Learning	7
2.2.3. Exploratory Learning	8
2.2.4. Active Learning	8
2.2.5. TEAL	10
2.2.6. MAL	11
2.3. Information Retrieval	12
2.3.1. Retrieval Process	13
2.3.2. Text Statistics	13
2.3.3. Term Weighting	15
2.4. Existing Solutions	18
2.5. Summary	24
3. Infret Prototype Version 1	27
3.1. Conceptual Design	27
3.1.1. Requirements	28
3.1.2. Conceptual Architecture	28
3.2. Development	29
3.2.1. Technology Decisions	31
3.2.2. Architecture	33
3.2.3. Features	33
3.3. Evaluation	40
3.3.1. Study Design	40
3.3.2. Settings and Instruments	40
3.3.3. Study participants	41
3.3.4. Findings and Discussions	41
3.4. Summary	48

Contents

4. Infret Prototype Version 2	49
4.1. Conceptual Design	49
4.1.1. Extended Conceptual Architecture	50
4.2. Development	50
4.2.1. Extended Architecture	52
4.2.2. Extended Features	52
4.3. Evaluation	60
4.3.1. Study Design	60
4.3.2. Settings and Instruments	61
4.3.3. Study participants	62
4.3.4. Findings and Discussions	66
4.4. Summary	79
5. Lessons Learned	81
6. Conclusion and Future Work	83
Bibliography	85
A. Infret Prototype Version 1 - Activities	91
B. Infret Prototype Version 1 - Survey	93
C. Infret Prototype Version 2 - Activities	97
C.1. Experienced Student Activites	97
C.2. Novice Student Activites	102
D. Infret Prototype Version 2 - Survey	107
E. UI Guideline	115

List of Figures

2.1.	Exploratory learning model	9
2.2.	General IR process	14
2.3.	First version of Infret	24
2.4.	Second version of Infret	25
3.1.	Infret prototype conceptual architecture	30
3.2.	Infret prototype architecture	34
3.3.	UI structure of the Infret prototype	35
3.4.	Collection selection modal	35
3.5.	Letter distribution	36
3.6.	Word distribution	37
3.7.	Word length distribution	37
3.8.	Word frequency rank	38
3.9.	Word plot rank	39
3.10.	Number of distinct words	39
3.11.	CES granular emotions for the Infret prototype	45
3.12.	CES four main emotions for the Infret prototype	47
4.1.	Infret's extended conceptual architecture	51
4.2.	The extended Infret architecture	53
4.3.	Infret prototype version 2 UI overview	54
4.4.	Cookie notification	54
4.5.	Tracking data structure	55
4.6.	Help tooltip shown when hovering over UI elements	56
4.7.	Help modal	57
4.8.	TF weight result example	58
4.9.	Modal displaying content of a collection document	59
4.10.	Example IDF results	59
4.11.	Example TF-IDF weight results	60
4.12.	Prior knowledge of novice students in 4 areas.	64
4.13.	Prior knowledge of experienced students in 4 areas.	65
4.14.	All CES emotions of experienced students for the second version of the Infret prototype	70
4.15.	All CES emotions of novice students for the second version of the Infret prototype	71
4.16.	CES four main emotions of experienced students for the second version of the Infret prototype	73

List of Figures

4.17. CES four main emotions of novice students for the second version of the Infret prototype	74
E.1. UI guideline 1	116
E.2. UI guideline 2	117

1. Introduction

This chapter is based on and content has been partially taken from (Bobić, Gütl, & Cheong, 2019). Common issues students face in class regularly have been identified by numerous studies (Dori, Hult, Breslow, & Belcher, 2007; D. Johnson, T. Johnson, & Smith, 1998). Three important issues which reduce the efficiency of the learning process are losing focus during longer lectures, having trouble understanding abstract concepts taught in some of the classes and retaining a low amount of knowledge after the class is finished.

An example of an important area where the understanding of abstract concepts is important is information retrieval (IR). Many concepts used in IR have their origins in text statistics and a multitude of mathematical concepts (Baeza-Yates, Ribeiro-Neto, et al., 1999). To make understanding of these concepts easier, students should be actively engaged during IR classes and provided with numerous exploration and interaction opportunities. By exploring and interacting with various concepts students can learn from their experiences and apply them to other concepts.

Experiential learning model describes such a learning process by defining learning as a continuous process where students gain experience, reflect on the gained experience and use the newly gained knowledge in their new experience (Kolb, 2014). However, this definition does not explicitly include the act of exploration. That is why exploratory learning expands the definition of learning with the introduction of the exploration concept (De Freitas & Neumann, 2009). To learn by gaining experience through exploration students have to be engaged. This is why the pedagogic approach active learning attempts to address the issue of focus loss in class by engaging students in a variety of activities which require active participation and not just silently sitting and listening to the lecture (Bonwell & Eison, 1991). Multiple evaluation studies indicate that this approach successfully increases students focus on lectures.

However, to address low knowledge retention and understanding of abstract concepts directly, new approaches were needed. One of these approaches is Technology-Enabled Active Learning (TEAL) which attempts to address low knowledge retention and abstract concept understanding by improving active learning with the introduction of interactive visualisations and many other concepts (Dori & Belcher, 2005). Another approach which expands on TEAL is Motivational Active Learning (MAL) which introduces gamification elements to motivate students in through the learning process and lower their resistance to the new engaging approach of learning (Pirker, Riffnaller-Schiefer, & Gütl,

1. Introduction

2014). MAL was used in an Information Search and Retrieval (ISR) class for many years where it was also evaluated.

1.1. Motivation and Focus

The experiences of using MAL in class indicate a need for an interactive experimentation and visualisation component to help students better understand the various abstract concepts taught in class. Therefore this thesis focuses on the conceptualisation, development and evaluation of a prototype and the second version of an interactive visual and exploratory tool called Infret for learning IR related concepts. The Infret prototype should enable students to explore text statistics concepts which are the basis for many other IR concepts in an interactive visual way. To do this a conceptual architecture and requirements will be defined and used to develop Infret. Once Infret is built it will be used by a group of students as part of a class activity. At the end of the activity, students should fill out an anonymous survey which will be later analysed, to see how Infret helps students learn and understand text statistics concepts. Based on these findings Infret will be further developed to also support more complex IR concepts. The second version of Infret will be used as part of a class activity in two distinct groups of students, where one will have more prior knowledge in IR and the other one less. To investigate how the second version of Infret helps students with different levels of prior knowledge in IR understand IR concepts, both groups of students will fill out an anonymous survey. Additionally, the feedback from the survey will be compared to the feedback gathered from the prototype. The data of this survey will be again carefully analysed to reach a conclusion.

1.2. Structure

This work is organised in the following way: Chapter 2 presents a background basis for this thesis. First, an overview of selected pedagogic approaches which attempt to address the aforementioned issues is made and the MAL approach is presented. This overview is followed by a brief introduction to the field of information retrieval and the relevant IR concepts for this thesis. Next selected existing solutions which are used as part of these approaches are presented and a conclusion to create a new solution is reached.

Chapter 3 starts by describing the conceptualisation of the Infret prototype which includes the definition of requirements, conceptual components and the definition of an overall conceptual architecture. Next, the final technology decisions are introduced, followed by Infrets final architecture and the implemented features. Finally, the goals of the study, the study settings and instruments and the participant demographics data are presented. The chapter concludes by

presenting and discussing the findings of the study.

Chapter 4 introduces the second version of Infret. First, the definition of new requirements and the extended conceptual architecture are briefly described. Next, the extended architecture and extended feature set are presented. The chapter concludes by introducing the evaluation details and discussing the evaluation findings.

Chapter 5 presents the lessons learned while working on this thesis from different views. These views include the literature survey, development and studies. Finally Chapter 6 concludes by summing up the completed work and findings in this thesis. Furthermore, this chapter also discusses possible future directions for Infret development.

2. Background and Related Work

Multiple studies (Dori et al., 2007; D. Johnson et al., 1998; Rowe, 1983; Ruhl, Hughes, & Schloss, 1987) have identified common issues students face while attending traditional face to face lectures. Students lose focus in a short amount of time during lectures and are unable to follow complex topics for longer periods. Some studies suggest students lose focus as soon as after 15 minutes (Prince, 2004) while others suggest students lose and gain focus in cycles and are more focused during interactive sections (questions or demonstrations) of a lecture (Bunce, Flens, & Neiles, 2010). Furthermore, they have trouble understanding abstract or intangible concepts usually taught at more advanced classes. Additionally, it was observed that many of them have low knowledge retention and have troubles remembering concepts taught after an extended period. To attempt to address these issues multiple new pedagogic approaches and tools have been created.

This chapter firstly presents some of the traditional pedagogic approaches. The learning models describing a more engaging way of teaching are presented next as the basis for new pedagogic approaches. Some of the popular interactive learning approaches that attempt to address the aforementioned issues students are facing are presented afterwards. Following is the introduction of the information retrieval (IR) field of study. Finally, the current software tools used for teaching IR and related fields are presented. The chapter concludes with the introduction of the tool whose development and evaluation will be the main focus of this thesis.

2.1. Basic Pedagogic Approaches

Just like every other field pedagogy experienced a technological revolution with the widespread adoption of Internet-connected devices (such as personal computers and smartphones). This led to the creation and use of many new teaching approaches. This section will present some of them. The most widely spread traditional teaching method is the face to face method which includes students sitting in a classroom and listening to the teacher's lecture. The teacher is usually in front of a whiteboard presenting the course content. In case students have any questions they raise their hand and ask the question once the teacher acknowledges them. This is a passive approach which assumes a one-way transmission of existing knowledge from lecturer to students. Distance

2. Background and Related Work

education unlike face to face teaching has students studying from wherever suits them best (Bates, 2005). The students are not required to sit in one classroom and do not need to listen to lectures in person. Today it heavily relies on the Internet to provide students with the course content, however, in the past it relied on the printed content. To describe learning where the World Wide Web is used to present the course content and to communicate with students in any form and intensity the term e-learning is used (Bates, 2005). Unlike face to face teaching, this is comprised of a lot of asynchronous communication through forums and also synchronous communication in the form of video calls or live chat. It should be noted, that the more the course relies on communication and information transfer over online solutions the more important the lecturer's online presence is (Henrich & Sieber, 2009). Blended learning has many advantages such as students choosing the location, time, pace and approach of studying (Alberts, Murray, Griffin, & Stephenson, 2007). However, the aforementioned advantages can also be disadvantages:

- The learning location and approach of studying the students choose can be inappropriate for studying and may hinder the students' study process.
- The freedom to choose the time of studying could lead to procrastination.
- The lack of a supportive social environment where students have to impose their deadlines might significantly slow down the learning process.

An alternative that represents a middle ground between traditional face-to-face learning and e-learning is blended learning. It integrates elements of face-to-face learning and e-learning in a carefully planned manner to leverage the strengths of both approaches (Alberts et al., 2007; Garrison & Kanuka, 2004). It is the lecturers' responsibility to decide the number of elements from each approach based on the teaching environment. Studies indicate blended learning is perceived positively, however, to facilitate a successful learning process lecturers should additionally try to foster a collaborative and interactive social educational environment in which students can be engaged (Güzer & Caner, 2014).

2.2. Interactive Learning

To attempt to tackle the difficulties students face while learning abstract concepts and gain their attention in class new pedagogic approaches were needed. These approaches count on interactivity and feedback during class as the factor that would help students learn more efficiently and interactively. This section firstly introduces supporting theories for the aforementioned interactive approaches such as the constructivist theory. Secondly, some of the aforementioned interactive approaches, starting with traditional and concluding with newer and more complex approaches which are built on the concepts of the

constructivist theory are introduced. The section concludes with the introduction of the Motivational Active Learning (MAL) approach which is the core approach for this thesis.

2.2.1. Constructivist theory

Based on the constructivist theory an individual creates new knowledge or understanding by combining their existing knowledge and past experiences with newly gained information or introduced concepts. In that way, every individual creates their way of understanding based on their past experiences. There are two main knowledge forming approaches constructivism describes. The first branch of constructivism is social constructivism which states that knowledge is built based on various environmental factors such as political views, ideological views, religious views and more. The second branch is psychological constructivism which states that knowledge or meaning is formed by the learner based on their existing body of knowledge. Additionally, the theory covers the knowledge forming process in social environments where individually constructed knowledge about a topic is discussed in a group. If the learners reach an agreement regarding the common definition of the individually formed knowledge about a phenomenon or in other words agree on the meaning for certain phenomena this meaning becomes formal knowledge (Richardson, 2003).

2.2.2. Experiential Learning

The experiential learning model is based on constructivism, however, it defines the learning process in a more precise manner. Experiential learning sets experience as the main actor in learning. It does so by combining three overlapping teaching models (Kolb, 2014).

The first model is the Lewinian model which describes a cyclic process which can be repeated multiple times to facilitate learning. The process starts with the student or learner experiencing some experience. The learner then collects data about the experience, observe it and analyses it. The analysis leads them to new implications which can be again explored and tested with the new experience. The process repeats multiple times. The emphasis is on using the experience to test our assumptions concluded from previously experienced events.

Dewey's model is similar to the Lewinian model, however, the emphasis is on being more explicit and describing learning as a feedback process. It describes learning as a process which transforms impulses into directed actions by observing and gaining knowledge. In other words, learning helps us transform our initial ideas and impulses driven by feelings into future planned actions which take into account possible consequences of themselves.

Finally, Piaget's model is similar to the two aforementioned models. This model

2. Background and Related Work

defines a similar outlook on learning by defining the key to learning as the combination of applying concepts to experiences in the real world and by extending or creating new concepts based on the experiences we gained.

In conclusion experiential learning defines learning as a neverending process of gaining experience which is transformed into new knowledge which can be transformed many times. Thus the process of gaining knowledge and adapting existing knowledge is emphasized. Knowledge is gained by a continuous process which happens multiple times. Each time the gained knowledge is transformed (Kolb, 2014).

2.2.3. Exploratory Learning

Exploratory learning which is depicted in Figure 2.3, is a learning model which expands the definition of experiential learning with the introduction of exploration and virtual environments (De Freitas & Neumann, 2009). Virtual environments enable students to explore and learn new concepts in their personalised way and their approach without the limitations of the real world. This personalised learning experience might lead to increased engagement and also to more self-motivated learning. Additionally, the students can receive feedback as soon as they act. Every exploration of the virtual environment can be a new experience for the students from which they can gain new knowledge. In this exploration-based approach to learning it is possible to provide students with various media representations of the same concept which enables them to study using their preferred media. Additionally, the exploration of virtual environments enables students to learn efficiently and engage in social interactions with others in the virtual world with whom they can share their ideas and discuss newly introduced concepts. The exploratory learning model is a multi-step cyclical learning model which starts with the exploration of a virtual or real-world which leads to experience. Students explore the world and learn from the new experience by performing various activities. They then reflect on the activities and experiences and attempt to integrate their newly gained knowledge into their existing collection of knowledge. The students test the newly gained knowledge in a virtual or real-world and gain again experience from it. A big emphasis is placed on learning being an open-ended process which continuously builds on existing knowledge and leverages problem-based and social interaction based approaches.

2.2.4. Active Learning

To facilitate learning by gaining experience through exploration, students need to be engaged in a variety of activities while attending a lecture and being introduced to new concepts. A widely used approach that leverages this type of lecturing is active learning. The term active learning was never precisely

The Exploratory Learning Model

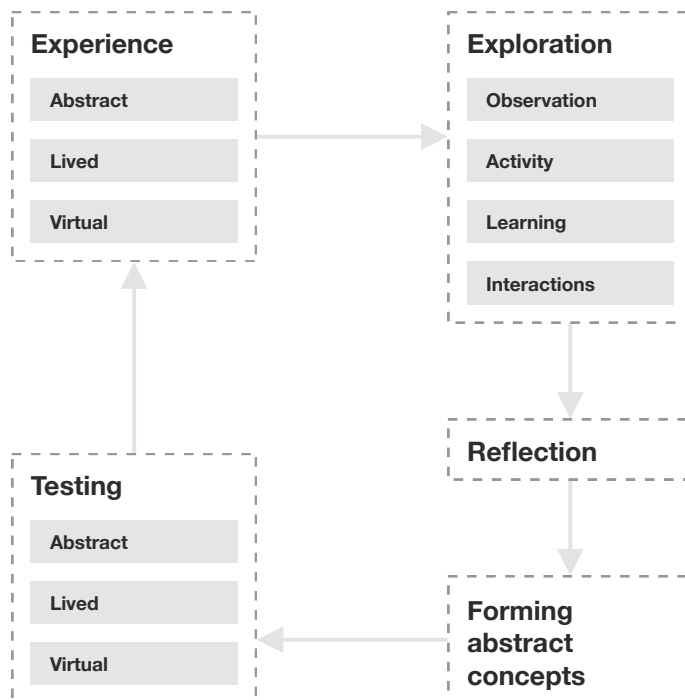


Figure 2.1.: The five steps of the exploratory learning model. Redrawn after (De Freitas & Neumann, 2009)

2. Background and Related Work

defined, however, Bonwell and Eison (1991) identified some of the properties of a class usually associated with active learning. These properties include:

- Students being more engaged and not just passively listening in class
- Emphasis is placed on developing student skills and not just on presenting information to them
- Students perform complex activities in class that include more than just passively listening to the lecture
- Larger focus on student exploration of their views

Based on the aforementioned properties it was suggested to define active learning as any learning process that *“involves students in doing things and thinking about the things they are doing”* (Bonwell & Eison, 1991). A more detailed definition is suggested in (Freeman et al., 2014) where active learning is described as a learning process which engages students through activities and discussion in class and puts greater weight on complex thinking and group work. The core elements of active learning are defined as introducing activities into traditional lectures and engaging students or promoting student engagement (Prince, 2004). An example of introducing activity during class is the introduction of short breaks in a lecture during which students work in pairs to clarify any misconceptions about the newly gained information. The evaluation of this approach indicates that students have better short and long term knowledge retention (Ruhl et al., 1987). The activity of having a 2-minute break to think about their notes and refine them encourages students to actively think about the newly learned concepts (Prince, 2004). An examination of 6000 student test data for an introductory physics course indicates that the interactive engagement methods drastically improved students performance in classes. Furthermore, an analysis of 255 studies of reported effects of active learning in science, technology, engineering, and mathematics (STEM) classes indicates that active learning helps raise average examination scores and that it is effective across all class sizes independent of how the lecturer implements the active learning activities (Freeman et al., 2014).

2.2.5. TEAL

Although active learning engages students and enables them to retain knowledge better it does not address the conceptual understanding of abstract concepts of students. In an attempt to address this issue Technology-Enabled Active Learning (TEAL) was introduced. This chapter introduces TEAL and is based on (Dori & Belcher, 2005) and (Dori et al., 2007) unless stated otherwise.

TEAL was developed in an attempt to involve first-year physics students in active learning with the aid of technology, help them visualise and improve their intuition about electromagnetic concepts and create a learning setting that encourages them to transform their ideas based on intuition into more

accepted scientific ideas. Multiple college physics curricula which emphasize the use of active learning motivated the development of TEAL. Additionally, a survey of 62 physics course indicates that the usage of interactive methods (such as collaborative learning) resulted in a substantial increase in conceptual understanding and problem-solving skills (Hake, 1998). The TEAL project was based on Studio Physic (Cummings, Marx, Thornton, & Kuhl, 1999) and SCALE-UP (Beichner & Saul, 2003) approaches.

TEAL combines short lectures with hands-on laboratory experiments, visualisations, problem-solving and discussions in between lectures in a technology-enhanced collaborative learning setting. Additionally, it integrates assessments through the course, which encouraged students to be prepared for new challenges and try to understand new concepts. Collaboration among students is facilitated by group work where they carry out hands-on experiments of exploratory nature using interactive visualisation tools, asking questions and discussing ideas. Short lectures can also be followed by multi-choice questions which are answered by each student using a personal response system. In case their answers are not unified the students discuss their answers again and answer once they conclude which answer is most probably correct. The experiments and explanations are usually aided by visualisations and animation to enable students to get a better understanding of the concepts and how they connect. It can be observed that TEAL's approaches and processes have their roots in constructivism and exploratory learning. Although TEAL was primarily designed for a physics course, it can be used as a model for science and tech courses in higher education.

The evaluation indicates that there was a substantial increase in conceptual understanding and a much lower failure rate compared to the more traditional approach of teaching. Students expressed positive and negative attitudes towards TEAL. Additionally, the study described in (Dori et al., 2007) which focused on the students who were participating in the TEAL course and the students who were part of the traditional course after one year to 18 months after the courses finished found that TEAL positively influenced students retention of electromagnetism concepts. Furthermore, students expressed resistance towards the new learning strategy as it required adjustment from the traditional way of teaching that they were used to.

2.2.6. MAL

An approach that builds on the TEAL approach and findings from its evaluation is Motivational Active Learning (MAL) created by Pirker et al. (2014). MAL combines the approaches used in TEAL with elements of gamification to motivate student learning. To address the needs of students with different learning approaches MAL uses a combination of approaches for teamwork and also for competition. The lectures are constructed out of short lectures,

2. Background and Related Work

accompanied by different engaging tasks, questions and assignments at the end of each short lecture. Some tasks require the students to work together in teams and solve problems. In addition, the students are given points (which do not influence their grades) based on their performance. The points in combination with leaderboards create a competitive environment (R. T. Johnson, Johnson, & Stanne, 1986). In case the students are not happy with their points they can re-submit the solution for selected activities. To get immediate feedback the results of the tasks, quizzes and assignments are submitted via an e-learning platform. This enables the teacher to track the students progress and at the same time provides students with immediate feedback regarding their newly acquired skills and knowledge. MAL was first tested at the Information Search and Retrieval (ISR) course in the winter semester of 2013 at the Graz University of Technology. The initial study included 28 students whose activities in the class were tracked. The results showed the students preferred working in small groups (2 students), their learning styles differed a lot and most students preferred points over grades and liked the possibility of re-submitting their solutions. The ranking of groups received mixed reactions from students as some of them liked the approach but others did not which further shows that students preference and learning styles are quite different.

2.3. Information Retrieval

To get a better understanding of the topic addressed in the MAL-based ISR class, this section provides an overview of the field of information retrieval (IR) and introduces the IR concepts relevant to this thesis. This section is mostly based on (Baeza-Yates, Ribeiro-Neto, et al., 1999).

IR is a field of study with a focus on information access, storage, organisation and presentation. It covers a broad spectre of topics such as system architecture, data visualisation, document categorisation, filtering and many more. One of the earliest uses of IR occurred with the use of tables of contents to navigate faster through books. Later, the need for IR increased with the introduction of the World Wide Web (WWW). The WWW enabled anyone to publish and read documents of any format on the web this introduced a need for more efficient ways of searching for new content on the WWW using IR techniques. To make the process of searching and parsing text documents easier each document can be represented as a set of keywords. To optimise the IR processes in large collections these text documents can be processed using various text operations such as stop-word¹ removal and stemming². These operations enable us to view a document collection as a set of index terms instead of a full set of keywords representing the entire document text. This representation and the intermediate

¹Common words appearing in a text which do not add information

²Reducing words to their grammatical root

representations can be then used by the IR system for further processing.

The main focus points of modern IR which need to be addressed are:

- Which approaches produce the best retrieval results?
- Which approaches will result in faster query processing?
- The quality of retrieval results is affected by users interaction with the system. How will understanding users interactions influence the design and retrieval results of IR systems?

2.3.1. Retrieval Process

To get a better understanding of a general IR process this sub-section describes a general retrieval process using the architecture shown in Figure 2.2.

To enable the system to carry out a retrieval process the database manager defines a text database by designating a document collection which will be used, the actions which will be performed on the documents and transform them into an internal representation and the document model (the structure of an individual document and the data that can be retrieved from it). Once everything is defined and the documents are converted to an internal representation the manager generates an index for the specified document collection, which is then used for the retrieval process and faster query evaluation in large document collections.

Once the internal representation of the document collection exists and the index is generated they can be used for a retrieval process. The user initiates the retrieval process by defining their need (for example with a sentence). The user need is then processed using the same text operations used to process the database and later might be processed also by query operations. This process produces a query which is used for searching the document collection with the help of the aforementioned index.

The resulting documents are ranked based on their relevance to the user need and returned to the user. In case the results are not relevant for the user, they can provide the IR system with relevance feedback by selecting the documents which are relevant for them. The system can then adjust the retrieval query based on the received feedback and return a new set of documents. This process can be repeated multiple times. However, in the real world users usually, just write a query which produces results. In case the users do not know how to formulate queries correctly they receive irrelevant results.

2.3.2. Text Statistics

IR includes many complex approaches and concepts which are used in the IR process, however, the basis for many of these concepts are text statistics. For example, the frequency of a word in a document and in the document collection

2. Background and Related Work

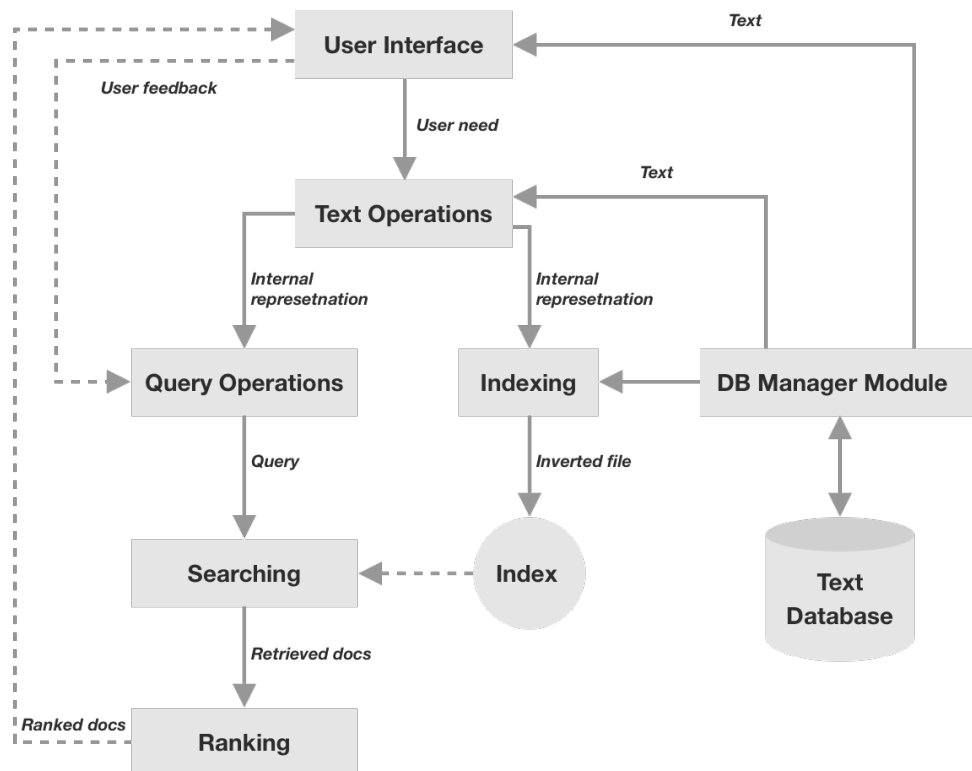


Figure 2.2.: Overview of a general architecture for an IR process. Taken and modified from (Baeza-Yates, Ribeiro-Neto, et al., 1999)

could be used to decide if a certain word would make a good index term to represent the document content (Baeza-Yates, Ribeiro-Neto, et al., 1999). If it appears in all documents it does not represent the topic of specific documents on the other hand if it appears often in only a few documents it might indicate the content of those documents. This is why this section presents selected text statistics which are later used as part of the thesis as standalone concepts and also as a basis for complex IR concepts.

Letter and Word Distribution The first text statistics are letter and word distribution which represent the number of each letter and word in a document or document collection. Even though these statistics are simple they can be used for various useful purposes such as detecting the language of the text, breaking simple cyphers, detecting stop words in text and more.

Word Length Distribution The word length distribution is another simple statistic which shows what are the most common word lengths and how often they occur in a document or document collection. Its value is calculated by counting the number of times a certain word of certain length appears.

Word Frequency Rank Unlike the word distribution which counts how many time a certain word repeats itself, the word frequency rank displays how many words are there with the same frequency in a document collection. The frequency with which most words appear has the highest rank.

Word Plot Rank Another statistic which deals with word frequencies is the word plot rank. It gives the highest rank to the word with the highest frequency, the second-highest rank with the second-highest frequency and so on.

Number of Distinct Words Finally, the number of distinct words describes how the number of words in the collection vocabulary³ changes with the growth of the number of all words in the collection.

2.3.3. Term Weighting

One of the basic concepts used in IR to represent the importance of words in text is term weighting. Term weighting is a process of applying weight values to terms in a document using various techniques. The assigned weight values indicate the relevance of a word for a specific document with various accuracy.

³Collection of unique words in a text document or collection of documents.

2. Background and Related Work

Term Frequency Weights The Term Frequency (TF) weights apply a numerical value to a term which is equal to its frequency in the document for which the weight is being calculated. The simplest TF weight is the binary TF weight equation (2.1). It applies a 1 as the weight value if the word exists in the document j and if not it applies a 0. This formula gives an overview of term occurrence in a document, however, it does not tell much more about the word relevance for a document as the values for all words in a document are 1.

$$tf_{i,j} = \begin{cases} 1, & \text{if } f_{i,j} > 0 \\ 0, & \text{otherwise} \end{cases} \quad (2.1)$$

However, past experiences indicate that the usage of complex weights which can express the intensity of the relevance of a word for a certain document instead of just binary values is more reliable when it comes to retrieval process (Salton & Buckley, 1988). That is why the TF weight equation (2.2) uses the term frequency as the weight value. In case the term does not appear in a document the weight is zero. The formula provides more information about the term compared to formula 2.1 as it gives an insight into how often the word appears in a certain document. However, this should not be taken as the final relevance measure for the document. Stop-words such as "a", "and", "I" tend to be the most used words in all documents and do not provide information about the document topic.

$$tf_{i,j} = \begin{cases} f_{i,j}, & \text{if } f_{i,j} > 0 \\ 0, & \text{otherwise} \end{cases} \quad (2.2)$$

The equation 2.3 normalises the frequency of a term i in document j ($f_{i,j}$) with the maximum term frequency of any term ($\max_i f_{i,j}$) in document j . These modifications help avoid large weight values in longer documents and provide more easily readable values for term weights. This leads to weight values between 0 and 1.

$$tf_{i,j} = \begin{cases} \frac{f_{i,j}}{\max_i f_{i,j}}, & \text{if } f_{i,j} > 0 \\ 0, & \text{otherwise} \end{cases} \quad (2.3)$$

An alternative approach to avoiding larger TF weight values is presented in equation 2.4, where (\log_2) of the term frequency is used in case the term frequency is larger than 0.

$$tf_{i,j} = \begin{cases} 1 + \log_2 f_{i,j}, & \text{if } f_{i,j} > 0 \\ 0, & \text{otherwise} \end{cases} \quad (2.4)$$

Finally equation 2.5 uses a similar approach as equation 2.3 but with an additional (K) parameter used for fine-tuning the weight values. This provides

the possibility to limit the weight values between the value (K) (which can be from 0 to 1) and 1.

$$tf_{i,j} = \begin{cases} K + (1 - K) \frac{f_{i,j}}{\max_i f_{i,j}}, & \text{if } f_{i,j} > 0 \\ 0, & \text{otherwise} \end{cases} \quad (2.5)$$

Inverse Document Frequency The aforementioned TF Weights show what term is most often represented in the document and set a weight value based on their TF, but, they consider all terms as equally important (Manning, Raghavan, & Schütze, 2010). However, some terms that occur often in a collection might not be as important or insightful for the documents in the collection as terms with fewer occurrences. To balance the high TF values a measure called the Inverse Document Frequency (IDF) is used. The value of IDF grows when the term occurs in a few documents and declines when the term occurs in many documents of a collection. Equation 2.6 us the basic IDF formula where the number of all documents (N) is divided by the number of documents the term (i) occurs in (n_i).

$$\log_2\left(\frac{N}{n_i}\right) \quad (2.6)$$

A term which occurs in all documents would receive an IDF value of 0 using equation 2.6. This might cause issues with complex weighting formulas that is why equation 2.7 adds a 1 to the ratio of all documents (N) and documents where the term (i) is present.

$$\log_2\left(1 + \frac{N}{n_i}\right) \quad (2.7)$$

Equation 2.8 is an alternative IDF formula which uses the largest number of documents ($\max_i n_i$) a word in a collection appears in as the numerator of the fraction inside of the logarithm.

$$\log_2\left(1 + \frac{\max_i n_i}{n_i}\right) \quad (2.8)$$

Another alternative for calculating the IDF is equation 2.9 which uses the ratio between the number of documents a word does not occur in ($N - n_i$) and the number of documents a word occurs in (n_i). This means the IDF will be negative for terms which occur in many documents and positive for words which occur in a few documents.

$$\log_2\left(\frac{N - n_i}{n_i}\right) \quad (2.9)$$

2. Background and Related Work

Term Frequency - Inverse Document Frequency Weights To maximise the efficiency of the retrieval process and produce weights which will accurately predict the relevance of words for certain documents a combination of both TF weights and IDF is needed (Salton & Buckley, 1988). Term Frequency - Inverse Document Frequency (TF-IDF) weights combine the concept of the TF weight with the IDF value to create a balanced weight value which will be high in case a term appears often in a small number of documents and low in case a term appears often in many documents. In case TF is 0 TF-IDF is also 0. A basic example of TF-IDF is equation 2.10 which combines the TF equation 2.2 with the IDF equation 2.6.

$$f_{i,j} * \log_2\left(\frac{N}{n_i}\right) \quad (2.10)$$

Another example of TF-IDF is equation 2.11 which combines equation 2.4 with the IDF equation 2.6. In this case there will be no zero values for TF-IDF if TF is 1.

$$(1 + \log_2 f_{i,j}) * \log_2\left(\frac{N}{n_i}\right) \quad (2.11)$$

Finally equation 2.12 combines equation 2.3 with the IDF equation 2.6.

$$\frac{f_{i,j}}{\max_i f_{i,j}} * \log_2\left(\frac{N}{n_i}\right) \quad (2.12)$$

2.4. Existing Solutions

With the introduction of approaches such as TEAL and MAL, interactivity and visualisations gained a central role in lectures. They give students the possibility to view and interact with the data in ways they could never do just by passively listening in class or doing basic homework assignments unless they dedicated their time to developing these tools themselves. With this, these tools also integrate exploratory learning concepts. To get an overview of what tools exist already this section focuses on a selected set of existing solutions in the field of IR. The field of IR is well established in today's society due to the high demand for fast and efficient ways of retrieving information. There are already quite a few educational solutions that attempt to address the issue of teaching IR to students interactively and intuitively. However, most of them focus on query performance, low-level concepts without the consideration for basic concepts that students might have trouble understanding.

Lucene Apache Lucene is a widely used opensource IR library developed using Java (Białecki, Muir, Ingersoll, & Imagination, 2012). It is being used in a

variety of commercial projects some of which have millions of users (such as Twitter, Netflix and Instagram). Some of the main features Lucene provides are:

- Conversion of documents into internal representations.
- Indexing of documents.
- Providing document statistics for various building blocks of document collections such as terms, documents and collections.
- Supporting a variety of query representations, manipulations on query results and providing a framework for supporting users with creation of custom queries.

The aforementioned list represents only a simplified collection of Lucene features. In addition to having a vast collection of features, Lucene is well optimised for speed and also optimal storage consumption and enables its users to choose between faster operation time or optimised storage usage.

Terrier Terrier is a publicly available IR framework (Ounis et al., 2005) for development of large scale IR applications. It supports indexing of document collections, document weighting and query expansion models. It scales based on the amount of provided data and works in a centralised or distributed way, depending on what is needed. It supports features such as tokenisation, removal of stop words, word stemming, calculating statistics about the document collection, automatic selection of document weighting model and retrieval approach and more. Furthermore, it is used in many real-world environments in universities and organisations. However, as López-García and Casheda (2011) already mentioned, Lucene and Terrier are too complex to be understood or changed by inexperienced students. Additionally they were not built with the MAL environment in mind.

IR Game The Information Retrieval Game (IR Game) was developed as a Web-based testing environment for query creation and evaluation using text and image test collections which were previously used only for IR experimentation and not for other practical purposes (Halttunen & Sormunen, 2000). The IR Game can be used to create a query, test the query on a textual or image document collection and visualise the evaluation results of the queries effectiveness. There are multiple forms of query performance feedback supported (for example precision/recall curves and highlighting of relevant documents in the results) which can be selected for each search task individually by the lecturer. IR Game additionally supports two modes of usage. First, the direct mode which allows the user to explore all areas of the IR Game and second, the exercise mode which provides the user access only to unlocked features. IR Game is constructed from 4 major components:

- A collection of search tasks for document retrieval

2. Background and Related Work

- Collections of correct search task results
- The frontend application for query formation and execution
- A feedback system for query performance analysis

The evaluation results indicate that most features depending on the student have either positive or negative effects on the learning process. For example, the authors explain that the feedback mechanism can be useful for students to understand how their queries perform it can be used to brute force the search task by changing the query until the feedback scores improve (for example increasing the precision/recall scores). Additionally, the authors concluded that the content of the text collection should be relevant for the students to spark their interest and improve motivation. Overall the IR Game offers an environment for experimenting with IR queries and evaluation them.

FIRS tool To improve students understanding of query weighting concepts and query evaluation in fuzzy IR systems⁴ (FIRS) a tool was created which provides a Web-based experimentation environment for students where they can analyse the performance of queries formulated in various ways (Lopez-Herrera et al., 2010). The students can formulate their queries with custom weights. The tool attempts to overcome the shortcomings of the IR Game (Halttunen & Sormunen, 2000) by visualising evaluation trees of weighted boolean queries, showing step by step results and enabling students to compare results of various queries and understand query semantics better. Additionally, the tool enables students to build custom test collections and analyse the performance of various weighted queries on them by using the precision and recall measurements. The evaluation results indicate that the tool helped students understand evaluation and query formulation related concepts and also that the intensity of use of the tool correlates with the average scores on exams, meaning the more they used the tool the better scores they got. However, it should be mentioned that the tool was evaluated on a small number of students (18) and that the correlation might not imply causation. Finally, the intensity was self-reported which could indicate that human bias was present.

IR-Toolbox The IR-Toolbox is a tool developed to attempt to help students learn the various processes of search engines without the need for programming knowledge (Efthimiadis & Freier, 2007). It enables students to explore document analysis and preprocessing, indexing, searching with the use of various weighting algorithms, query formulation and evaluation of results. It is built using Lucene⁵ as the search engine. To the knowledge of the thesis author, the authors of IR-Toolbox did not provide any further implementation or evaluation and the IR-Toolbox webpage is not accessible.

⁴IR systems which use fuzzy IR techniques.

⁵<https://lucene.apache.org/>

VIRLab VIRLab is a Web-based IR tool created to help the development and studying of IR models by enabling users to easily implement new retrieval functions, evaluate them and analyse their performance (Fang, Wu, Yang, & Zhai, 2014). It is aimed at researchers and students with limited programming experience. VIRLab provides an interactive interface that enables the implementation of a retrieval function, the evaluation of retrieval functions over multiple document collections and result analysis through search engines or pair-wise comparisons of results. The users can write the code for a retrieval function which combines various provided collection statistics inside the interface. Multiple IR system instances can be configured by selecting the desired retrieval function and document collections. Queries can be written using custom code or selected from a predefined list for a specific collection. VIRLab also enables users to make side by side comparisons of two search engines. Additionally, VIRLab provides the option to evaluate a user created retrieval function over a set of document collections automatically. Predefined document collections are provided with official query relevance judgments which are displayed for easier evaluation of custom functions. The 10 most effective retrieval functions for each collection are displayed in a leaderboard to promote the experimental study of IR. The users can also use the aforementioned comparison feature to compare their retrieval functions to the retrieval functions on the leaderboard.

Xtrieval Web Lab Xtrieval Web Lab is another Web-based application which enables users to learn about the IR process (Wilhelm-Stein, Kahl, & Eibl, 2017). It is created for non-programmers as the main user group and structured from multiple configurable IR components which can be arranged and evaluated in a variety of ways with real-world data. Additionally, it leverages gamification to motivate the user to learn. Xtrieval Web Lab displays a preview of the processed data to the users to help them configure the components and understand how each component manipulates data. Xtrieval Web Lab then processes the data on the server and returns the results to the user interface (UI) where they can be displayed using a variety of visualisations. Additionally, there is a big emphasis on customizability by enabling users to customise the list of stop words, configure components, reorder components and more. The gamification aspects of Xtrieval Web Lab are realised by a level system which enables the users to level up and create more complex retrieval systems until all components are unlocked. Furthermore, to motivate students to explore additional components and concepts a system of achievements and badges was implemented. Gaining achievements also means gaining access to new components. It was shown that the students achieved better results when motivated by this method. Students are introduced to the system using assignments which require them to use more components and new parts of the UI over time. Finally, it provides an exploration mode where students have access to all components to facilitate exploration.

2. Background and Related Work

IR-Base The IR-Base is a proposed modular open-source IR framework for the creation of new IR components, applications and for learning IR concepts (Calado, Cardoso-Cachopo, & Oliveira, 2007). It would come with pre-implemented Java class skeletons and guidelines on how to use them to create an IR component which can be used by others. The target group of users are researchers, professionals, lecturers and students. It could be used for prototyping and for learning IR concepts. The main goal of the framework is to have a wide variety of applications and to not be limited to only specific concepts. Each components internal functioning should be completely independent of other components but, should still be able to communicate to any other component. An interesting idea that the authors mention is the component pool which would serve as a repository of components from where anyone could search for and use already implemented components. Additionally, the community could add new components to the repository. Although IR-Base is a great idea it not implemented to the knowledge of the thesis author.

IR-Components IR-Components is a Java framework which attempts to help teachers avoid creating IR tools from scratch by providing already partially integrated components in form of interfaces, fully built classes, algorithm skeletons and Application programming interfaces (API) which can be extended by students (López-García & CACHED, 2011). Each component has an implemented version which can be reimplemented by the end-user and used instead of the already implemented component. By providing partially built components it offers its users modularity and customizability while enabling them to easily interconnect various components. The authors suggest that the framework has multiple usages such as quickly creating an IR application by using the already built modules, rebuilding some of the modules and combining them with the already built modules or using the already built modules to explain how a certain component works. Additionally, they provide three different educational applications ideas:

- The MiniCrawler which would be an implementation of a web crawler which downloads a targeted web pages data in a specific data structure.
- The MiniIndexer whose main purpose would be to create a document and term index for a document collection also supports multiple ways of storing these indexes and some term processing approaches like stop word removal.
- The MiniSearcher which would be meant to be an IR system which can be accessed through multiple input methods and configured to use a variety of weighting mechanisms and term processing approaches.

Each of the aforementioned application ideas was designed to provide additional educational value to the students such as teaching the students how a

crawling and indexing algorithms work, why some characters could cause problems when implementing an indexer, familiarise themselves with the process of resolving user queries and more. The evaluation indicates that the students think IR-components were helpful in their exercises. However, it should be noted that the feedback survey and evaluation was carried out with six students. The authors additionally reference IR-Base (Calado et al., 2007) but argue that it is a simpler solution than IR-Components. Although both IR-base and IR-components provide good solutions for building complex solutions they are limited to one programming language which limits the use to students proficient in that specific language. An interesting idea to explore would be a solution that would enable students to develop their components in multiple languages and use them in a container application.

ISRApp Despite the existence of different tools and projects to improve information retrieval teaching, there was a need to create a tool named ISRApp specifically for a MAL based ISR course at Graz University of Technology which would be tailored to the course (Ziessler, 2018). ISRApp is a Java-based tool which enables students to select a document collection, view its text statistics properties in form of graphs, use a web scraper to create a custom document collection and a simple search engine. It additionally supports custom implementations of some components of ISRApp such as the search engine and the web crawler. The evaluation results indicate that the ISRApp helped students better understand text statistics and made students feel positive emotions while using the application.

Infret Even though the ISRApp received positive feedback it was limited only to desktop computers and could not be easily updated remotely if needed. Based on the results and feedback received from the ISRApp and the survey of some selected existing solutions, a prototype of a new Web-based tool called Infret was created as an interactive component of MAL (Bobić, Gütl, & Cheong, 2019). It attempts to help students understand the concepts of IR using visualisations and interaction in the same ISR class as the ISRApp. Infret is accessible from any Internet-connected device such as a tablet or a personal computer. It enabled students to select a text collection or a slice of a text collection, view various text statistics related to the selected collection and control the visualisations. A screenshot of Infrets UI can be seen in Figure 2.3. Infret was used in the ISR class as a central part of a text statistics exercise where students had to solve a variety of exercises using Infret. The evaluation results indicate that Infrets usability is above average and that the students generally liked working with it. However, students also suggested many potential improvements which were taken into account for further development.

The second version of Infret was built based on the requirements and feedback from the first version of Infret (Bobić, Cheong, Filippou, Cheong, & Gütl, 2019).

2. Background and Related Work

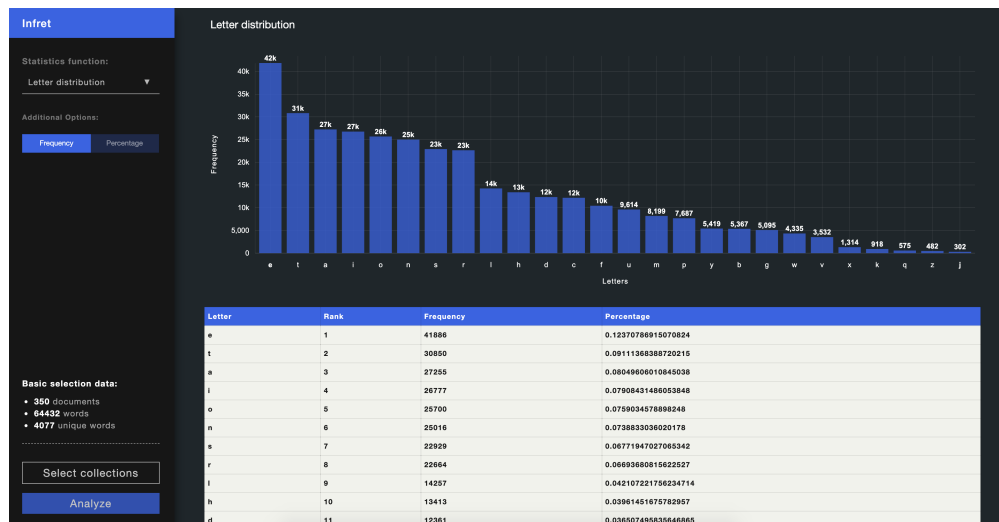


Figure 2.3.: First version of Infret displaying the word distribution of a text collection

The second version includes many UI enhancements, new term weighting concepts and visualisations for them and new interaction tracking capabilities. It enables students to switch between text statistics and term weighting and to choose among three term weighing concepts and various formulas for each concept. A screenshot of the updated UI and a new heatmap visualisation can be seen in Figure 2.4. The second version of Infret was used in an ISR class and in a database design and development class with two groups of students with various levels of experience. A novice group which had little previous experience with IR concepts and an experienced group which was at the end of an ISR class.

2.5. Summary

This chapter provided an overview of older as well as newer interactive pedagogic approaches based on concepts of constructivism and exploration which attempt to enable students to learn more efficiently, have better knowledge retention and focus for longer periods while being interactively engaged in the class. One of the latest approaches (MAL) focuses on bringing the interactivity and exploration of TEAL to an ISR class while helping students accept the new interactive approach with the use of gamification. To provide one of the central components of MAL which is interactive visualisation and exploration to an ISR class, there is a need for an interactive visualisation and exploration tool. A set of selected existing tools have been presented in this chapter, however, they either focus on low-level implementations of IR concepts, on query performance analysis or are simply too complex to use in a class. Additionally, many of the presented tools do not have an emphasis on visualising abstract

2.5. Summary



Figure 2.4.: second version of Infret displaying the term frequency weights of terms in a document collection.

concepts present in IR and do not focus on presenting the basic concepts of IR. Furthermore, some of the existing solutions are either large and complex frameworks which are meant to be used for development and not so much for learning or are just theoretical concepts. Due to these limitations, a prototype of an IR visualisation and exploration tool called Infret has been developed and evaluated. The following chapters of this thesis will focus on the process of creation, expansion and evaluation of Infret with new features which can also be found in (Bobić, Gütl, & Cheong, 2019) and (Bobić, Cheong, et al., 2019).

3. Infret Prototype Version 1

The concept of an educational exploratory and experimental IR application named Infret was created based on the findings and ideas from the literature review, the requirements of MAL (Pirker et al., 2014) and the findings from the ISRAApp (Ziessler, 2018). Infret aims to serve as a visual exploratory and interactive component for a MAL based ISR class and attempt to help students learn IR concepts in an interactive and exploratory manner.

This chapter is structured as follows. Section 3.1 describes the conceptual design, defines the main requirements and introduces the conceptual architecture of the prototype. Section 3.2 introduces the final technology decisions, the architecture and the features of the Infret prototype. Finally, the evaluation setting, environment and findings are presented in Section 3.3. Parts of this content were also described in (Bobić, Gütl, & Cheong, 2019).

3.1. Conceptual Design

Infret should provide an exploratory and experiential environment by using various visualisations and interactions to present a variety of IR concepts. Concepts which could be presented are text statistics, term weighting, natural language processing (NLP) approaches and more. To display these concepts Infret should include at least some text-based document collections. Additionally, it should enable users to create custom document collections and utilise web crawling (as seen in the ISRAApp (Ziessler, 2018) and other reviewed solutions) for the creation of custom document collections. To get a better insight and understanding of various collections users should be able to compare properties such as text statistics of multiple collections. Infret could provide a personalised learning experience by enabling students to code custom IR concepts inside of Infret (for example custom term weighting functions, text statistics or crawlers). Additionally, Infret should be platform agnostic¹ to avoid stress during setup on a users device and widen the potential user base for Infret. Infret is an educational tool and should, therefore, be integrated with various learning management systems (LMS) such as Moodle². Furthermore, it should support the creation of learning tasks, questions based on exercises and grading of tasks with the help of an LMS to provide a seamless experience for the students and the lecturer. To

¹ Accessible from any computer, tablet or phone and any operating system

² <https://moodle.org/>

3. Infret Prototype Version 1

motivate students Infret could offer a levelling system which would guide them through various exercises and components of Infret. Finally, the lecturer should have an overview of students progress and interaction data with Infret in the form of a password-protected administration panel. It could help the lecturer to identify components in which students are either particularly interested or have issues understanding. This section is based on and content has been partially taken from (Bobić, Gütl, & Cheong, 2019).

3.1.1. Requirements

Based on the aforementioned concept a set of high-level requirements which serve as guidelines for the development of the Infret prototype was defined:

Functional Requirements

- To attempt to help students better understand abstract IR concepts, the data relating to these concepts should be presented in a visually interactive way.
- The first Infret prototype should attempt to help students understand and familiarise themselves with text statistics (which are also the basis for many other IR concepts) of document collections.
- Infret should support text-based document collections.
- To be accessible by a large number of users, Infret should be platform agnostic.

Non-functional Requirements

- Infret should be developed with MAL in mind and provide students with possibilities to explore and engage with various IR concepts.
- As Infret might be developed by multiple people at the same time in the future its architecture should be modular and support development on multiple areas at the same time.
- Technologies used for building Infret should be welcoming to new developers.
- Infret's architecture should support the addition of potential extensions.

3.1.2. Conceptual Architecture

The conceptual architecture seen in Figure 3.1 was designed based on the aforementioned requirements. Each block represents a component. Communication between components and the data passed from one component to another is displayed with an arrow signalling the flow of communication and an italic text

describing the sent data. Possible further expansions are marked with three dots.

Infret's conceptual architecture is divided into separate independent components to enable multiple developers to work on different features at the same time.

To support text-based document collections Infret uses the pre-processing document collection component to read a collection using the file reader, select an appropriate parser based on the file format using the parser selector and transform (parse) the file to an internal representation format with the help of file parser. The newly formatted data is then stored in the data storage component as an internal representation.

Infret implements a processing component which provides various text statistics functions to help users understand and familiarise themselves with text statistics. It provides the text statistics functions by reading the internal representation data, calculating one of the six supported statistics on the data and storing the calculated text statistics values in the data storage as cached statistics.

The side panel is used by the user to chose a collection and a statistics function they would like to run on the chosen collection. The cached statistics are then read from the data storage and sent to the client. To help users understand text statistics, Infret displays the statistics in a variety of interactive visual components such as bar charts, line charts or tables. Finally, the user can adjust the visualisations with filters in the side panel. The Infret prototype offers the following filters:

- Change graph orientation, which is used to change the graph orientation in case there are two graphs available for a statistic. The user can choose between having graphs side by side for easier comparison or having them one under the other for detailed investigation.
- The frequency/percentage switch enables the user to choose between viewing the values in a raw frequency format or percentage format.
- To enable the users to change the number of items displayed in a chart, Infret's side panel displays an interactive slider element.

3.2. Development

To develop an efficiently designed Web-based application which enables multiple developers to implement separate features at the same time and new developers to join easily the technologies for Infret had to be chosen carefully. These technology choices are presented in Section 3.2.1. Once the technologies were chosen the architecture which is described in Section 3.2.2 was defined based on the conceptual architecture. Finally, the features implemented in the

3. Infret Prototype Version 1

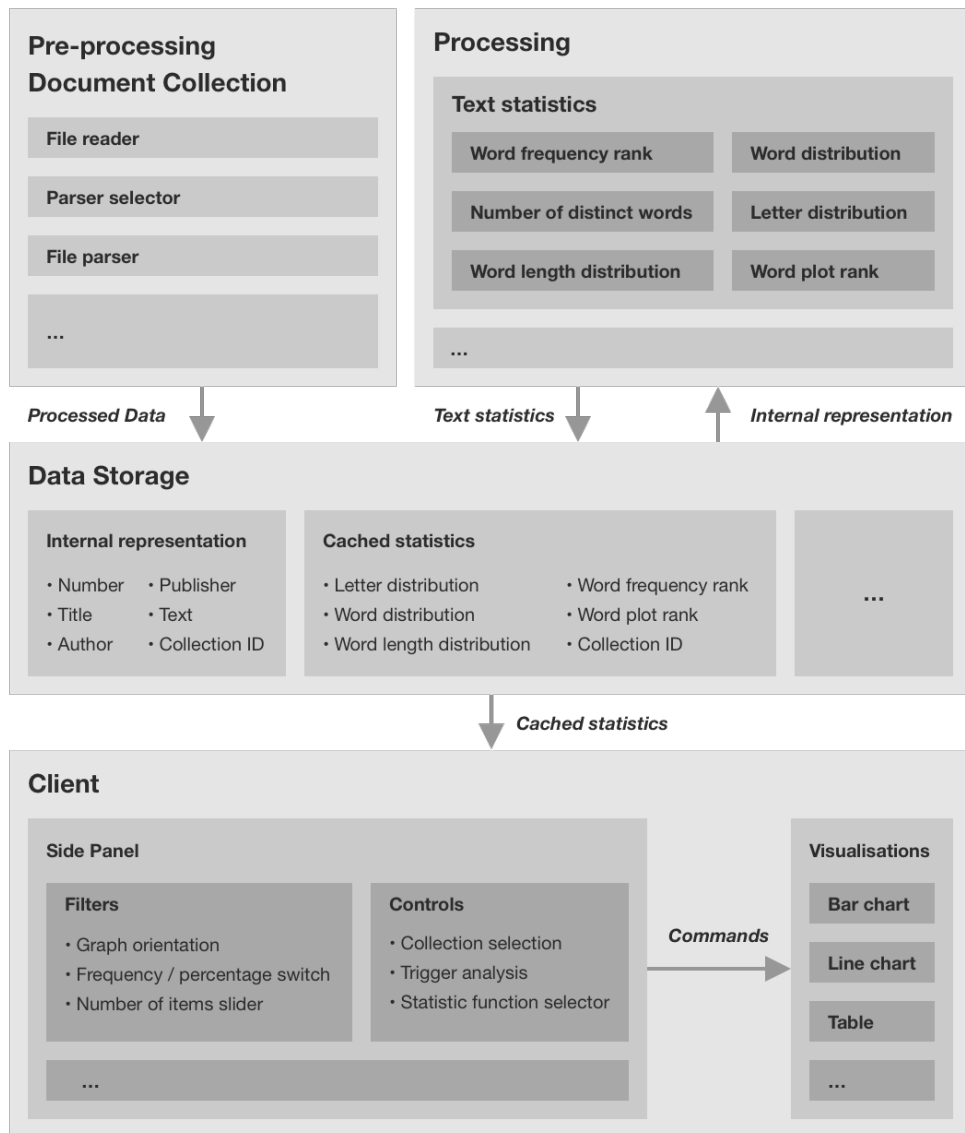


Figure 3.1.: The conceptual architecture of the Infret prototype. Individual components are displayed in grey blocks. The arrows represent the data flow from one component to another and the italic text next to the arrows describes the type of data being sent. Taken and modified from (Bobić, Gütl, & Cheong, 2019).

Infret prototype are presented in Section 3.2.3. This section is based on and content has been partially taken from (Bobić, Gütl, & Cheong, 2019).

3.2.1. Technology Decisions

To develop Infret efficiently and provide developers with the tools and environment needed to expand Infret in the future the technologies for building Infret had to be chosen carefully. The requirements and conceptual architecture defined in Section 3.1 were considered while choosing the technologies.

Client Side

To provide Infret to a wide variety of users and make it platform agnostic, Infret runs in a Web browser. Because of that, the only suitable language for the client-side is JavaScript. However, writing the client-side in pure JavaScript would make it much harder to design a modular architecture. This is why the most actively developed and supported client-side frameworks at the time of the development (2017) were considered. The two JavaScript frameworks which offered the most support for developers were extensible and enabled a modular-based architecture were Angular³ version 2 and React⁴ version 15. Additionally, both frameworks are welcoming to new developers by being well documented. However, only Angular enforces a strict coding style and has a code style guide which helps developers write code in a structured manner. Additionally, Angular supports TypeScript⁵ as its main language. TypeScript is a superset language of JavaScript which transpiles into JavaScript and brings a typing system and new language features that are not yet supported by the current version of JavaScript. This enables new developers who might have only experience with JavaScript to get easily adjusted to new features and the typing system while simultaneously helping them avoid bugs, resolve existing bugs and document their code. Due to the aforementioned benefits TypeScript is the language of choice for the client-side.

To enable students to explore text statistics concepts in a visually interactive way, NVD3.js⁶ was chosen as the visualisation library for providing various interactive visualisations on the client-side. It provides the necessary customisation options and a wide range of visualisations needed for presenting various statistics.

³<https://angular.io/>

⁴<https://reactjs.org/>

⁵<https://www.typescriptlang.org/>

⁶<http://nvd3.org/>

3. Infret Prototype Version 1

Server Side

To attempt to support new developers in adjusting to the code base, Infret's client- and server-side should be based on similar technologies. This is why the server-side is built using version 6 of the JavaScript runtime Node.js⁷. Node.js is based on the C language and offers access to many system features which are not available with plain JavaScript.

Like the client-side, the server-side requires a framework to guide its structure in an organised and modular manner. Hapi.js⁸ version 15 was chosen to achieve the modularity and extensibility of the server-side. It is a well-structured framework built by developers at Walmart to handle increased visitor traffic during special events like Black Friday⁹.

To store the internal representations of various text statistics Infret also has to include a database. The document-oriented no-SQL database MongoDB¹⁰ and the widely used relational database MySQL¹¹ were considered. Both have their downsides as the relational MySQL is not ideal for document representation and the document-oriented MongoDB had issues with loss of data at the time of this comparison. The final decision to use MySQL was motivated by the fact that it was a safe and proven way of storing data and that it provided an efficient way of retrieving the needed data. To use the MySQL database in Node.js an Object-Relational Mapper called Sequelize¹² was used.

Other Development Tools

To enable easier development for larger groups of developers as well as new developers multiple additional tools are used.

Webpack¹³ is used for various task automation such as transpiling TypeScript and SASS into JavaScript and CSS respectively. Additionally, it is used to switch between the development and production version of Infret. The production version uses the correct access data for the server and the code is minified to provide faster loading times.

Furthermore linters such as stylelint¹⁴, ESLint¹⁵ and TSLint¹⁶ are used to aid developers in writing consistently formatted code which is essential for larger projects and developer teams.

⁷<https://nodejs.org/en/>

⁸<https://hapijs.com/>

⁹A shopping holiday in the US on the Friday after Thanksgiving, marking the start of the Christmas shopping season

¹⁰<https://www.mongodb.com/>

¹¹<https://www.mysql.com/>

¹²<https://sequelize.org/>

¹³<https://webpack.js.org/>

¹⁴<https://stylelint.io/>

¹⁵<https://eslint.org/>

¹⁶<https://palantir.github.io/tslint/>

Finally Git¹⁷ is used as the versioning system of choice to enable multiple developers to work on the project at the same time seamlessly and to have a better overview of feature development and the state of Infret.

3.2.2. Architecture

Infret was built based on the aforementioned technology decisions and conceptual architecture. The final architecture of the Infret prototype can be seen in Figure 3.2. Each block represents a component group. The lines between various components indicate these components are interconnected and have one or multiple connections where various data is transferred in one or both directions. Due to the number of connections between components the depicted architecture is simplified and does not represent all connections between components. The data transmitted through these connections is described with italic text.

3.2.3. Features

The Infret prototype provides users with a simplistic UI seen in Figure 3.3. It is constructed from a side panel (A) which enables the users to manipulate the visualisations and data panel (B) which displays the analysis results with various visualisations. The user opens the collection selection module seen in Figure 3.4 by pressing the “Select collections” button (C). Once they select the collections they would like to analyse they can see the details of the selected collection at the bottom of the side panel (E). Finally, they can select one of the provided statistics functions (D) and trigger an analysis (C). The analysis results are visualised in the data panel (B).

Collection selection

The provided collection selection modal seen in Figure 3.4 enables users to select one of the items on the list of all collections (F). When a collection is selected they are shown the total number of documents in the selected collections (G). Finally, the user completes the selection by pressing the “Finish selection” button (H)

Letter Distribution The results of the letter distribution static can be seen in Figure 3.5. It presents the user with a distribution of all letters in the text collection with a bar chart. The results are additionally presented in a table for easier use with other software tools. The user may view the distribution of letters either as raw frequencies or as a percentage.

¹⁷<https://git-scm.com/>

3. Infret Prototype Version 1



Figure 3.2.: The final architecture of the Infret prototype. Blocks represent component groups. Connections between components are represented with lines between components and represent more than one connection going in one or both directions. The arrows represent the direction of the data flow, while the italic text describes the data transmitted through the connections.

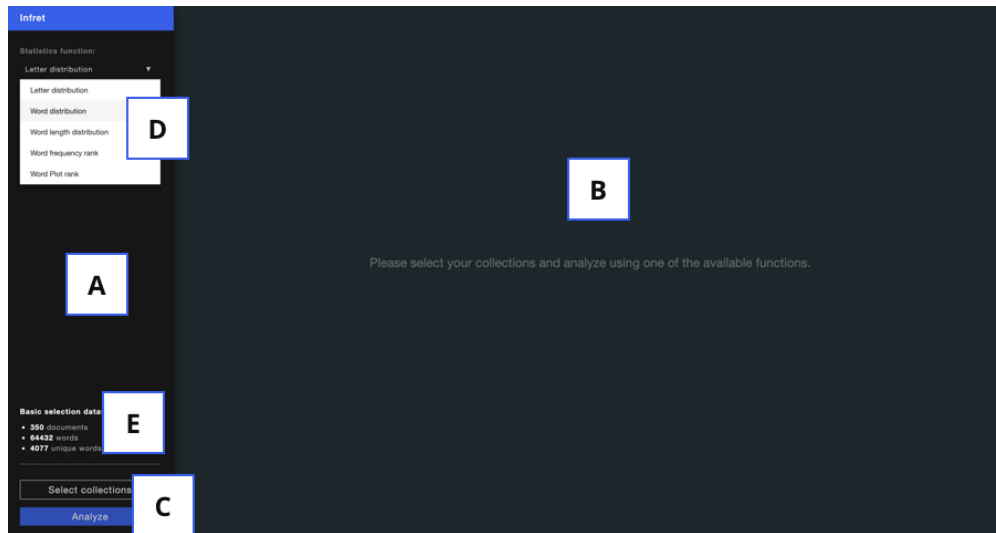


Figure 3.3.: UI structure of the Infret prototype

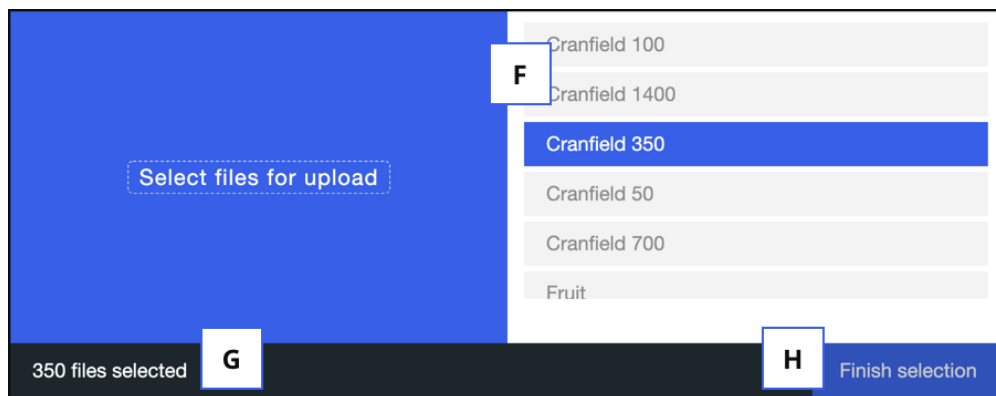


Figure 3.4.: Collection selection modal

3. Infret Prototype Version 1

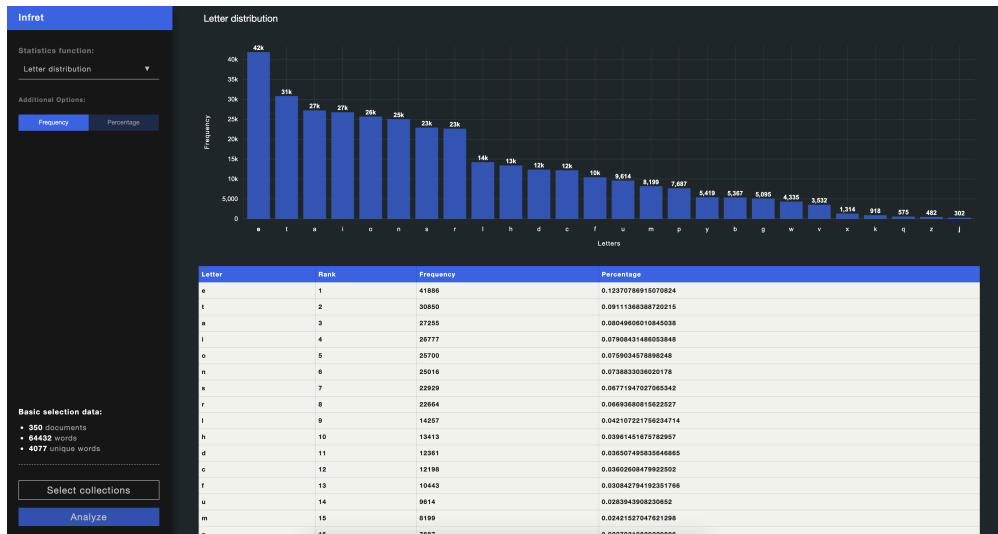


Figure 3.5.: Letter distribution

Word Distribution The word distribution results pictured in Figure 3.6 provide the user with information about the statistical distribution of words in the selected text collection. As with the letter distribution the user can switch between viewing raw frequencies and percentages. Additionally, they can choose to have the two charts positioned side by side and tune the number of elements they would like to view. The bar chart presents the word distribution where the left-most word has the highest frequency and the right-most word has the lowest frequency. To provide the users with a smooth experience the number of words shown in the graph is limited to 250. The line chart, on the other hand, presents the word rank on the x-axis and probability * rank on the y-axis.

Word Length Distribution The word length distribution displayed in Figure 3.7 provides the user with the statistics about the word lengths in the document collection. The results are displayed with a bar chart and a table where the x-axis of the chart displays the word length while the y-axis displays the frequency or number of words with that length. The user has an option to switch between frequency and percentage.

Word Frequency Rank The word frequency rank is seen in Figure 3.8 describes the rank of a given word frequency or in other words what are the most and least common word frequencies. The results are presented with a bar chart where the x-axis presents the word frequency in ascending order and the y-axis presents the number of words with that frequency. The user can lower the number of items in the chart for easier investigation.

3.2. Development

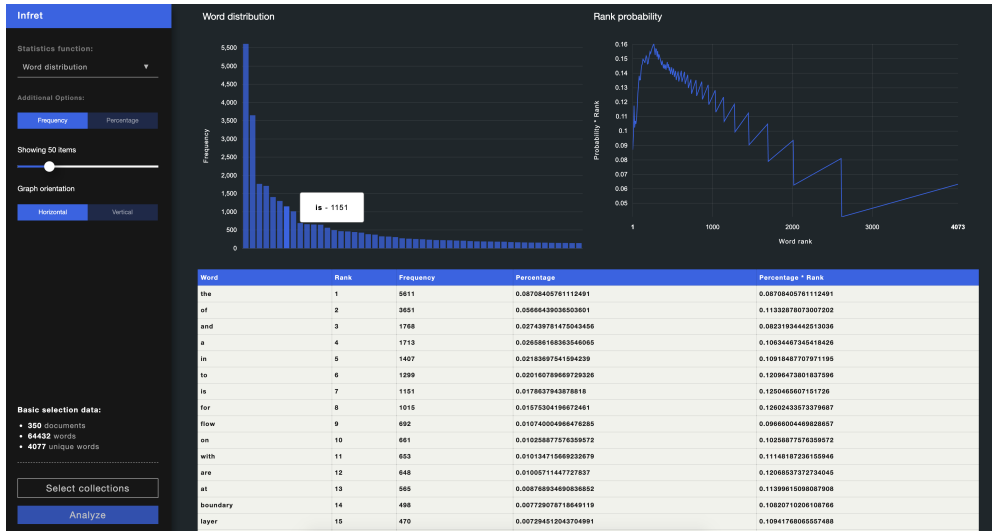


Figure 3.6.: Word distribution

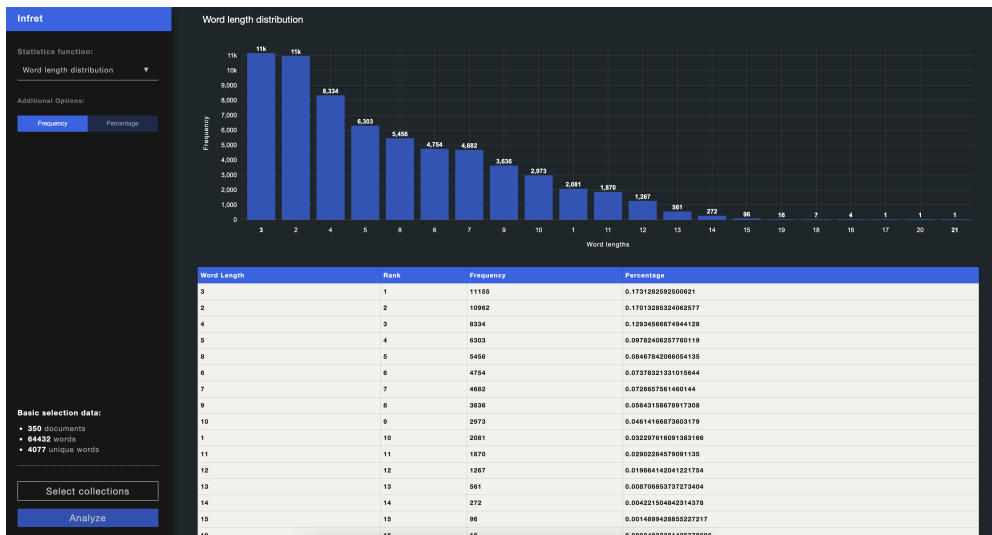


Figure 3.7.: Word length distribution

3. Infret Prototype Version 1

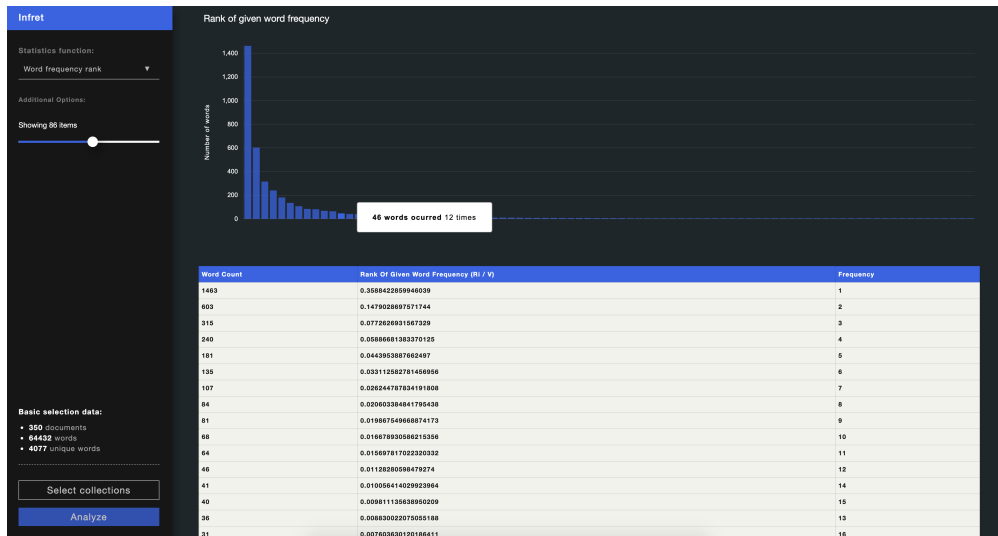


Figure 3.8.: Word frequency rank

Word Plot Rank The word plot rank displayed in Figure 3.9 compares the word frequency in a logarithmic scale on the y-axis of the line chart with the word rank on a logarithmic scale on the x-axis of the chart.

Number of distinct words The number of distinct words pictured in Figure 3.10 provides an insight in how the number of unique words in the vocabulary of a collection (y-axis of line chart) grows with the number of words in the same collection (x-axis of line chart) over time. It can be noticed that the number of words in vocabulary grows slower than words in the collection.

3.2. Development

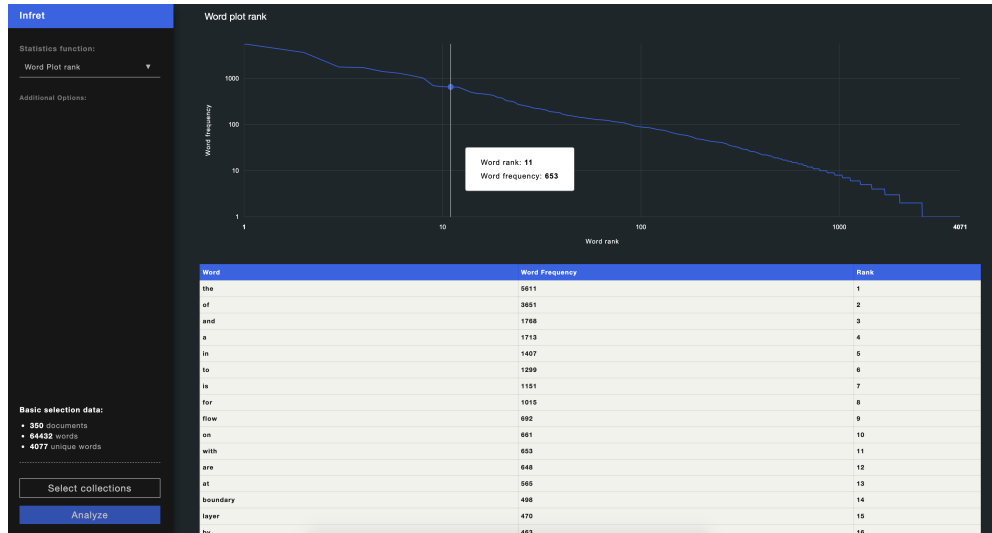


Figure 3.9.: Word plot rank



Figure 3.10.: Number of distinct words

3. Infret Prototype Version 1

3.3. Evaluation

Once Infret was developed it was used as part of a student class activity. Feedback data was gathered using a survey and later analysed to reach final evaluation conclusions. This section provides a detailed overview of the Infret evaluation process including the study design, instruments and environment, the demographics of the study participants and finally the findings and conclusions. This section is based on and content has been partially taken from (Bobić, Gütl, & Cheong, 2019).

3.3.1. Study Design

The primary goal of this study was to understand how Infret helps students learn text statistics concepts. Furthermore, the study should reveal what is the overall students experience of the first version of the Infret prototype and can Infret support students in understanding or deepening their understanding of text statistics concepts. Students were provided with a set of activities in the form of questions and instructions which guided them through various text statistic related concepts. These activities can be seen in Appendix A. The activities can be completed by navigating through Infret to acquire the data and by processing the data using knowledge learned in the ISR course which they were attending. The students had one week to complete the activities. Once the activities were finalised and all questions answered the students submitted their solutions. At the end of the deadline, students were provided with a link to an anonymous online survey seen in Appendix B, which they filled out. The survey results were then analysed using excel and python to answer the aforementioned research questions.

3.3.2. Settings and Instruments

Students were provided with five activities seen in Appendix A which enabled them to explore various text statistics concepts. The first activity instructs the students to explore a text collection and give an overview of their findings. The next activity consists of investigating the probability distribution of single words using the letter distribution functionality in Infret. The third activity required the students to explore the word distribution of first 40 words. The word distribution is explored in the fourth activity and the number of distinct words is explored in the fifth activity. Activities three to five require additional use of knowledge gained during the ISR class. The students were provided with links to an anonymous online survey once the submission deadline was over. The provided survey seen in Appendix B was split into multiple question groups based on the topic they were associated with. The first set of question gathers various demographics about the participants, such as the age, gender

and study program. The next set consists of ten Likert-scale questions from the standardised System Usability Scale (Brooke et al., 1996) (SUS) used to evaluate the overall usability of a tool. Based on the answers of these questions it is possible to calculate a usability score which can then be compared to previous versions of Infret or other similar tools to get an understanding of how the usability of the evaluated version of Infret compares to the usability of other tools. Furthermore, the survey includes a set of twelve questions from the Computer Emotion Scale (Kay & Loverock, 2008) (CES) used to measure four emotions of students while using Infret: happiness, sadness, anxiety and anger. The survey is concluded with a collection of feedback questions used for the general improvement of Infret. Once all of the results were gathered the data was exported to an excel file and analysed using built-in functions of Excel or by using Python.

3.3.3. Study participants

The study participants were 25 master students from the Graz University of Technology participating in the ISR course in the winter semester of the school year 2017/2018. The students were taught various IR concepts during the semester and were provided with the activities and the survey at the end of the semester when they already had some experience with IR.

The majority of the students were enrolled in the study program “Computer Science” or “Software Development and Business Management”. 22 (88%) of the participants were male and 3 (12%) were female. Furthermore, 15 (60%) of the students were between 20 and 25 years old, 9 (36%) were between 26 and 30 years old and 1 (4%) student was between 31 and 40 years old.

3.3.4. Findings and Discussions

The majority (88%) of surveyed students agreed they like experiments and hands-on activities. When asked if experiments and hands-on activities help them better understand the theory and methods 22 (88%) of the students agreed. Detailed answers to the aforementioned questions can be seen in Table 3.1. The results received a year before in a similar setting for the ISRApp resemble the aforementioned answers received for Infret. This could indicate that students generally feel that interactive tools and experiments help them learn better and such tools should be included as part of the curriculum.

When asked about their experience with Infret, the majority (56%) agreed Infret helped them better understand and reflect on aspects of text statistics and 16 (64%) students also indicated they would like to use such tool for other subjects in ISR. Students mentioned term weighting, image data retrieval and digital libraries as possible areas of expansion for Infret. Detailed answer statistics for the questions on Infret can be seen in Table 3.2.

3. Infret Prototype Version 1

Table 3.1.: Students opinion about experiments and hands-on activities

	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
Q1	7 (28%)	15 (60%)	1 (4%)	2 (8%)	0 (0%)
Q2	12 (48%)	10 (40%)	2 (8%)	1 (4%)	0 (0%)

Answers to the following survey statements: Q1 - I like experiments and hands-on activities; Q2 - Generally, experiments and hands-on activities help me to better understand theory and methods

Table 3.2.: Students feedback about Infret

	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
S1	2 (8%)	12 (48%)	7 (28%)	1 (4%)	3 (12%)
S2	2 (8%)	14 (56%)	4 (16%)	3 (12%)	2 (8%)

Answers to the following survey statements: S1 - This tool has helped me to better understand and reflect aspects on text statistics; S2 - I would like to use such tool for other subjects in ISR.

The answers in Table 3.2 indicate that Infret successfully helps students understand concepts of text statistics and most would like to use it in other areas of ISR. When asked what they liked about the tool students mentioned the visualisations, interactivity, the design and animations:

“I liked the design of the tool and that you could hover over the graphs and get additional information depending on the item you are investigating.”

“The simplicity of the user interface and the representation of the data”

“I liked the design a lot. The usage of the software was also very intuitive. Everything was structured well and easy to understand.”

“The animated graphs and easy navigation were quite good.”

Even though the aforementioned feedback implies students were pleased with Infret, possible future improvements were also identified. When asked what they disliked about Infret students mentioned they had interaction issues with the slider element, lack of configuration options for the visualisations, issues selecting columns of a table and animations:

“That you could not select single columns for copying the values. Aside from that, there was not anything else that I did not like at all.”

“The slider to change the size of the (for example) words. Is hard to tod [sic] with the mouse, it is easier to type in the number or use the arrow keys for it. It is odd the first time to click analyze again after you choose in the drop-down menu a method”

“I didn’t like that you couldn’t alter the graph’s representation more. The only parameter you could change was the number of items showing.”

Additionally, students were asked for improvement suggestions. They suggested adding an option to view the content of the document in the document collection, improve the layout on smaller devices, filter the results in tables, adding more customisation to visualisations and more:

“Maybe that you can check the collection itself.”

“ With a smaller device (such as my laptop) it was sometimes a bit frustrating to scroll in the tool because there were 2 scrollbars. and when I open the tool I can’t directly see the analyze button”

“seeing a list of documents from the collection, possibility to add/remove columns from analyzed results, customizable graphs (label, color, aspectratio), add possibility to change the data of the collection to experiment a little with it”

“Formulars (describe variables better, etc. . .)”

“Maybe providing functionality to further investigate the data, e.g. sorting the tables and items according to different attributes.”

These answers indicate there are a lot of potential improvements which could be addressed in the future and used to offer a better experience to Infret users. To inspect how Infret’s usability compares to the usability of the ISRApp the standardised SUS (Brooke et al., 1996) survey consisting of 10 Likert scale questions was used. The results indicate Infret’s mean usability score is 76.9 and its standard deviation is 11.8. The average usability score is 68, which means Infret’s usability is above average (Brooke, 2013). Compared to the SUS score of the ISRApp Infret’s score is slightly better, however, that could be explained by the lack of a search engine module which would add to the complexity of the UI. The detailed SUS score of both Infret and the ISRApp can be seen in Table 3.3.

Table 3.3.: SUS Scores for the Infret prototype and the ISRApp

	SUS Mean	SUS Standard deviation
Infret	76.9	11.8
ISRApp	73.8	11.1

Finally, to get a better understanding of how Infret affects students while using it the standardised CES (Kay & Loverock, 2008) survey was used. The

3. Infret Prototype Version 1

survey consists of 12 questions used to measure 4 main distinct emotions and 12 sub-emotions. Each of the 12 CES questions ask the participants how often they experience a certain emotion while using Infret. The participants can answer using one of the following answers:

1. None of the time
2. Some of the time
3. Most of the time
4. All of the time

The mean value and standard deviation for all emotions can be seen in Table 3.4. Furthermore, a box plot displaying the distribution of answers can be seen in Figure 3.11. Most negative emotions have a mean value close to 1 which indicates students experienced them rarely while using Infret. However, it can be seen that some students reported feeling dispirited and disheartened some of the time. On the other hand, positive emotions such as satisfaction and curiosity have values between 2 and 3 which shows that many students experience those emotions some of the time or even most of the time. Furthermore, 75% of the students felt excited while using Infret at least some of the time.

3.3. Evaluation

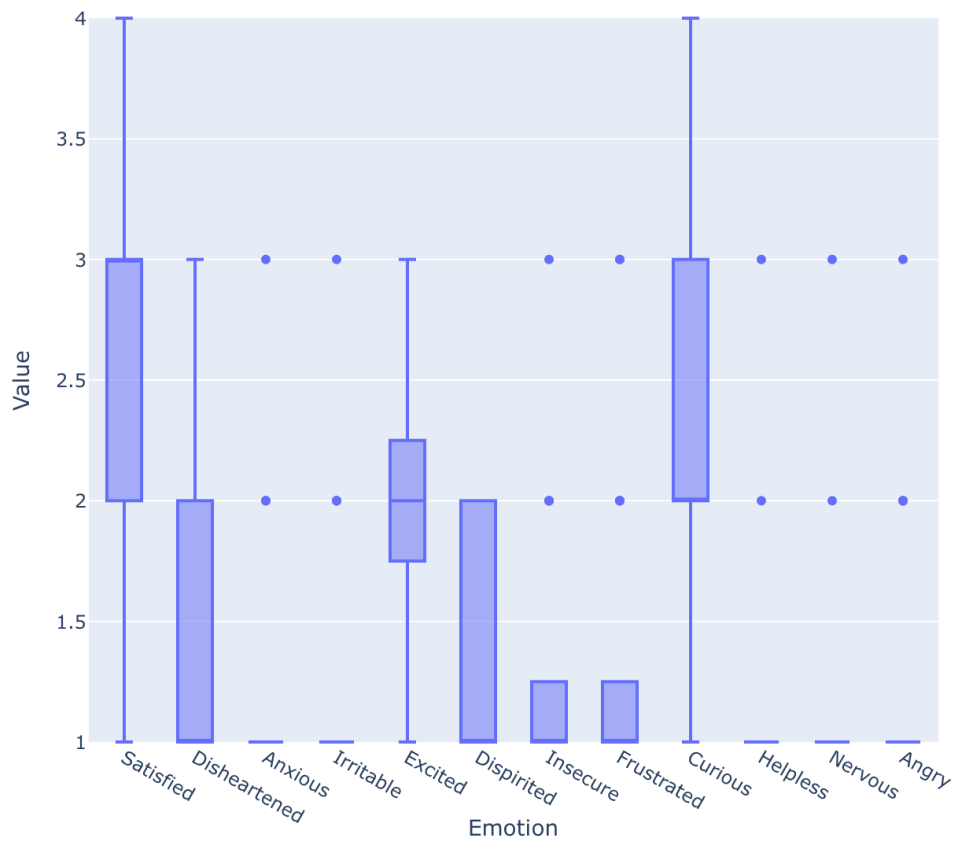


Figure 3.11.: CES granular emotions for the Infret prototype

3. Infret Prototype Version 1

Table 3.4.: CES statistics for granular emotions

Emotion	Mean	Standard deviation
Satisfied	2.6	0.71
Disheartened	1.4	0.58
Anxious	1.2	0.5
Irritable	1.24	0.52
Excited	2	0.71
Dispirited	1.28	0.46
Insecure	1.28	0.54
Frustrated	1.32	0.63
Curious	2.36	0.76
Helpless	1.12	0.44
Nervous	1.12	0.44
Angry	1.24	0.52

The distribution of the values for the four main emotions calculated from all of the measured emotions can be seen in Figure 3.12 and the mean and standard deviation values are displayed in Table 3.5. Just like in the aforementioned statistics for all granular emotions it can be seen that happiness is the predominant emotion students experience. This indicates that Infret mostly causes students to experience positive emotions and motivates them to be curious about IR concepts. However, it should be noted some of the students experienced also negative emotions such as sadness or anger.

3.3. Evaluation

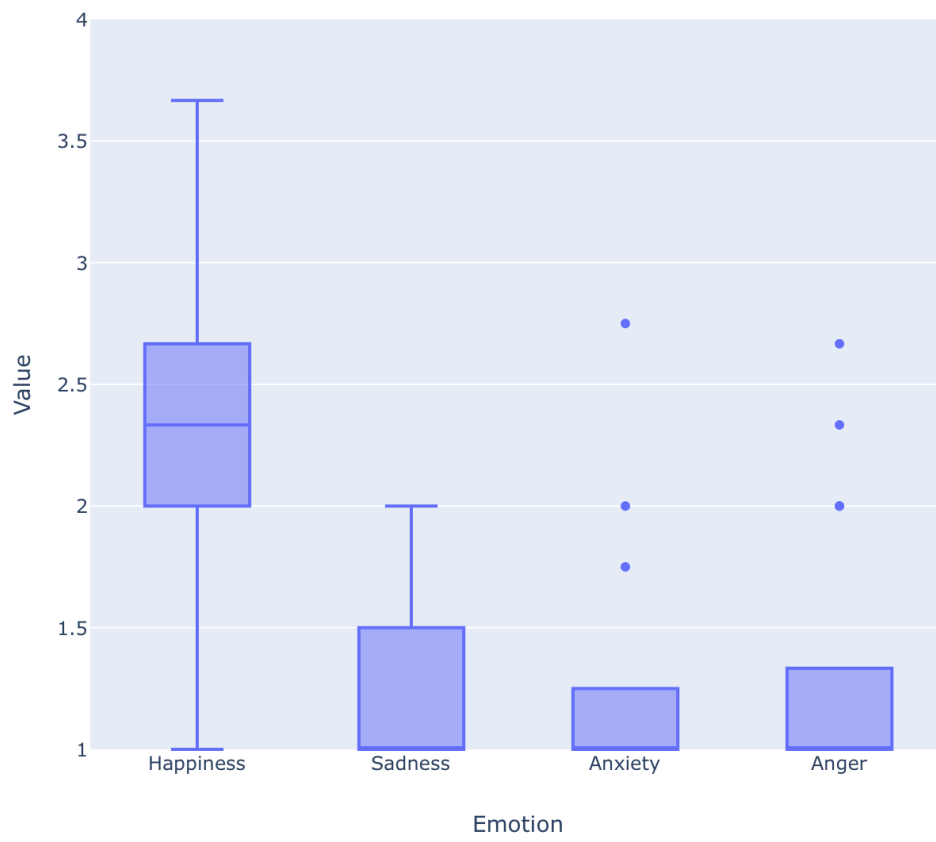


Figure 3.12.: CES four main emotions for the Infret prototype

3. Infret Prototype Version 1

Table 3.5.: CES statistics for four main emotions

Emotion	Mean	Standard deviation
Happiness	2.32	0.54
Sadness	1.34	0.40
Anxiety	1.18	0.41
Anger	1.27	0.47

3.4. Summary

The prototype of a Web-based interactive exploratory IR tool called Infret was conceptualised, developed and evaluated based on the findings of the literature survey and the feedback received during the evaluation of the ISRAApp(Ziessler, 2018). It was created to serve as a visual exploratory component in a MAL-based ISR class. Students can use Infret to select one of the predefined text document collections and investigate one of the 6 implemented text statistics functions. The statistics are presented in a visual way using various graphs and also in tabular form for use in other applications. Users can customise the visualisations to some extent using custom controls in the UI of Infret. Students used Infret at the end of the semester as the last activity at the ISR course. They filled out an anonymous online survey once they finished the activities. The survey data indicates Infret's usability score is slightly higher than the usability score of the ISRAApp and the average usability score. Additionally, students mostly experienced positive emotions while using Infret. Students mostly liked Infret and would like to use it for other aspects of ISR, however, they also mentioned potential Infret improvements which would make the experience better.

4. Infret Prototype Version 2

A new version of the Infret prototype was conceptualised, developed and evaluated based on the evaluation findings of the first version (Bobić, Gütl, & Cheong, 2019). The second version aims to help students understand term weighting concepts in addition to text statistics.

This chapter is structured as follows. Section 4.1 introduces the requirements for the second version of Infret and the extended conceptual architecture. The final extended architecture and new features are introduced in Section 4.2. Finally, the evaluation process, environment and results are introduced and interpreted in Section 4.3. Parts of this chapter are based on and content has been partially taken from (Bobić, Cheong, et al., 2019).

4.1. Conceptual Design

The second version of Infret is built based on the evaluation results and findings of the prototype and the initial concept (Bobić, Cheong, et al., 2019). This section is based on and content has been partially taken from (Bobić, Cheong, et al., 2019). To expand the conceptual architecture of Infret a set of extended requirements was defined based on the primary concept:

Functional Requirements

- Infret should enable users to explore and interact with multiple term weighting concepts.
- The term weighting concepts should be provided in an interactive visual manner.
- To get a better insight into how users use Infret and interact with it, Infret should track user activity.
- Infret should attempt to help users with various levels of prior IR knowledge by providing help modals. The modals should display theoretical explanations of IR concepts implemented in Infret.
- The users should be able to view and explore document content of text documents used in Infret.
- Interaction with the side panel should be improved by removing bugs and adding UI elements which could help users interact with Infret.
- Infret's UI should be improved to accommodate smaller screens.

4. Infret Prototype Version 2

4.1.1. Extended Conceptual Architecture

The conceptual architecture seen in Figure 4.1 was designed based on the aforementioned requirements and is taken from (Bobić, Cheong, et al., 2019). Each block represents a component. The light grey blocks represent components implemented in the first version of the Infret prototype while the dark grey and black blocks represent the components introduced in the second version of Infret. Communication between components and the data passed from one component to another is displayed with an arrow signalling the flow of data and an italic text describing the sent data. Possible further expansions are marked with three dots.

The processing component was expanded with various weighting functions to enable users exploration of various term weighting concepts. The processing component reads the internal representation from the data storage, calculates selected term weighting concepts and stores the results back in the data storage. To provide the term weighting concepts interactively and visually, the client reads the term weighting data from the data storage and displays it using one of the available visualisations. The client was updated with a heatmap visualisation to intuitively display the Term Frequency (TF) weights and Term Frequency - Inverse Document Frequency (TF-IDF) weights.

The expanded client provides the users with new controls. The first new control is an activity selector which enables users to select the concept (activity) they would like to explore. They can choose between text statistics and term weights. The second new control is the weighting formula selector which enables users to select one of the formulas of the selected term weighting concept. The redesigned client additionally provides new filters for tuning parameters of certain term weighting formulas and the removal of stop words in results.

To provide users with help modals Infret implements a help component in the client. The help component displays tooltips with details about the functionality of various UI elements. Additionally, it displays the help modal which provides theoretical explanations for selected weighting function.

To gain a better insight into how users interact with Infret the new version includes a tracking component. Each time a user interacts with Infret the interaction information (tracking data) is sent to the tracking component. The data is parsed by the component in an appropriate format and stored in the data storage in the tracking table. The data can later be retrieved and analysed using external tools.

4.2. Development

Once the requirements for the second version of the Infret prototype were defined and the conceptual architecture was expanded the development could start. As there was no need for new technology introduction, this section

4.2. Development

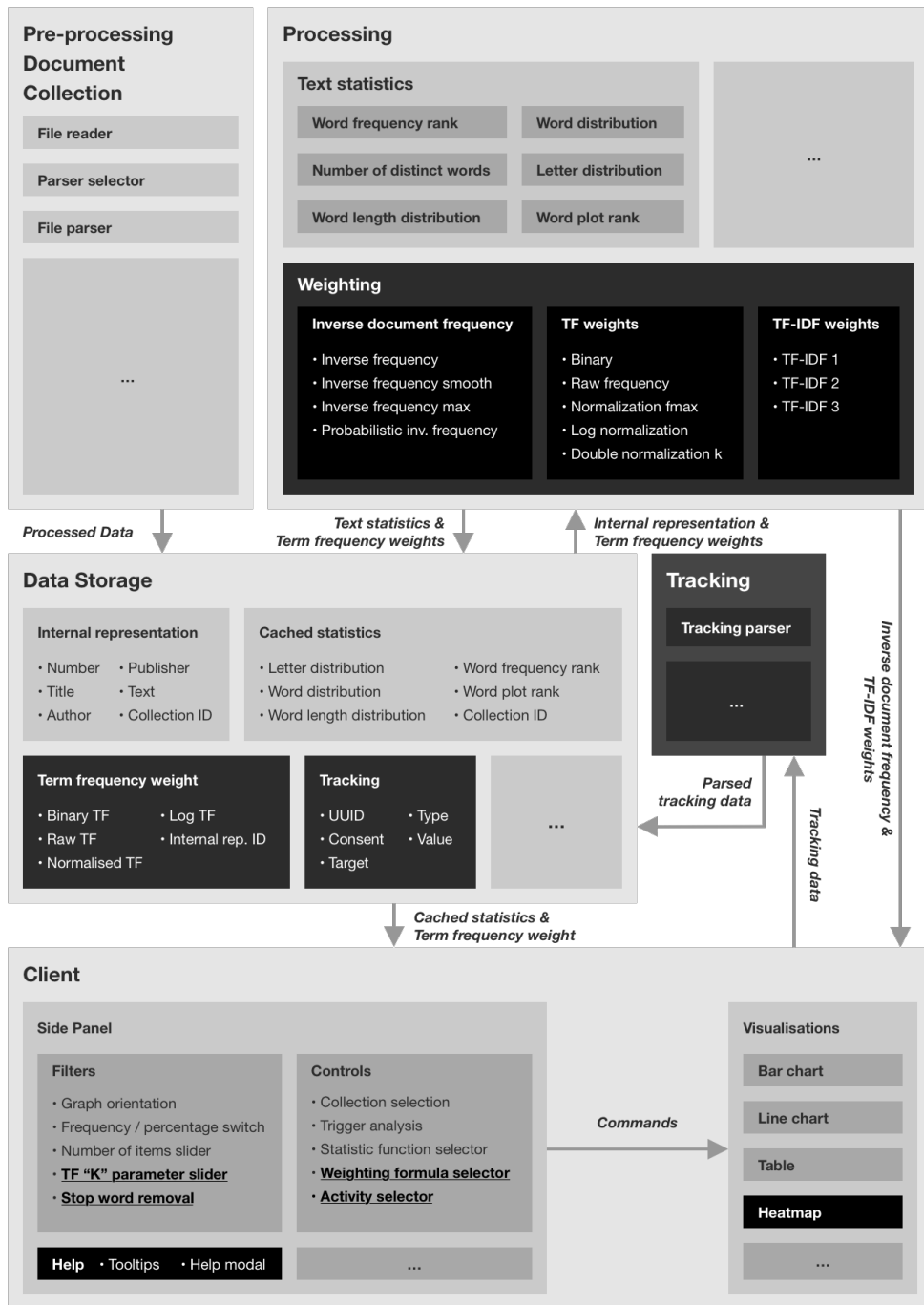


Figure 4.1: Infret's extended conceptual architecture. The newly added components are displayed in black or dark grey blocks. The connections between them are depicted with arrow lines and the data sent through them is described with italic text. Taken and adapted after (Bobić, Cheong, Filippou, Cheong, & Gütl, 2019)

4. Infret Prototype Version 2

presents the expanded architecture and the newly introduced features of the second version. This section is based on and content has been partially taken from (Bobić, Cheong, et al., 2019).

4.2.1. Extended Architecture

The second version of Infret is designed and built based on the extended requirements and extended conceptual architecture. The extended architecture is depicted in Figure 4.2. Black and dark grey blocks represent newly introduced components. Lines between components represent multiple connections. The italic text next to the arrow lines describes the data sent through these connections.

4.2.2. Extended Features

The second version of the infret prototype presents users with an extended and improved UI displayed in Figure 4.3. The new UI was modified based on the feedback received from the first prototype evaluation. Therefore, the new UIs sidebar embraces a top-down workflow¹. Furthermore, the new UI enables users to explore additional IR concepts in the area of term weighting. Once a user selects a collection they would like to work with, they chose an activity domain with the newly introduced activity selector (A). Next, the user either selects a weighting or statistic concept which they would like to explore with the concept selection dropdown (B). Additionally, they can choose one of the multiple formula implementations found under the formulae selection dropdown (D). Finally, the “Analyze” button can be clicked which will trigger the analysis and provide the user with results in the data panel.

Interaction tracking The first new feature of the second version of Infret is the ability to track user interaction with the UI anonymously and provide valuable feedback for further research and improvements to Infret. A new user is presented with a modal displayed in Figure 4.4 asking for permission to use the anonymously collected data for research purposes. Furthermore, it informs the user of what data is collected. The collected data includes the identifier, consent answer, time of interaction and the action itself. A detailed structure of the tracked data is depicted in Figure 4.5. Once they choose to accept or deny the request a new unique identifier number is generated and displayed in the lower-left corner of the application. This identifier can be later used to connect the users’ interaction data with their survey answers in an anonymous manner.

¹Top level controls such as selection of collections are positioned at the top, while lower-level controls such as the customisation of views are positioned at the bottom.

4.2. Development

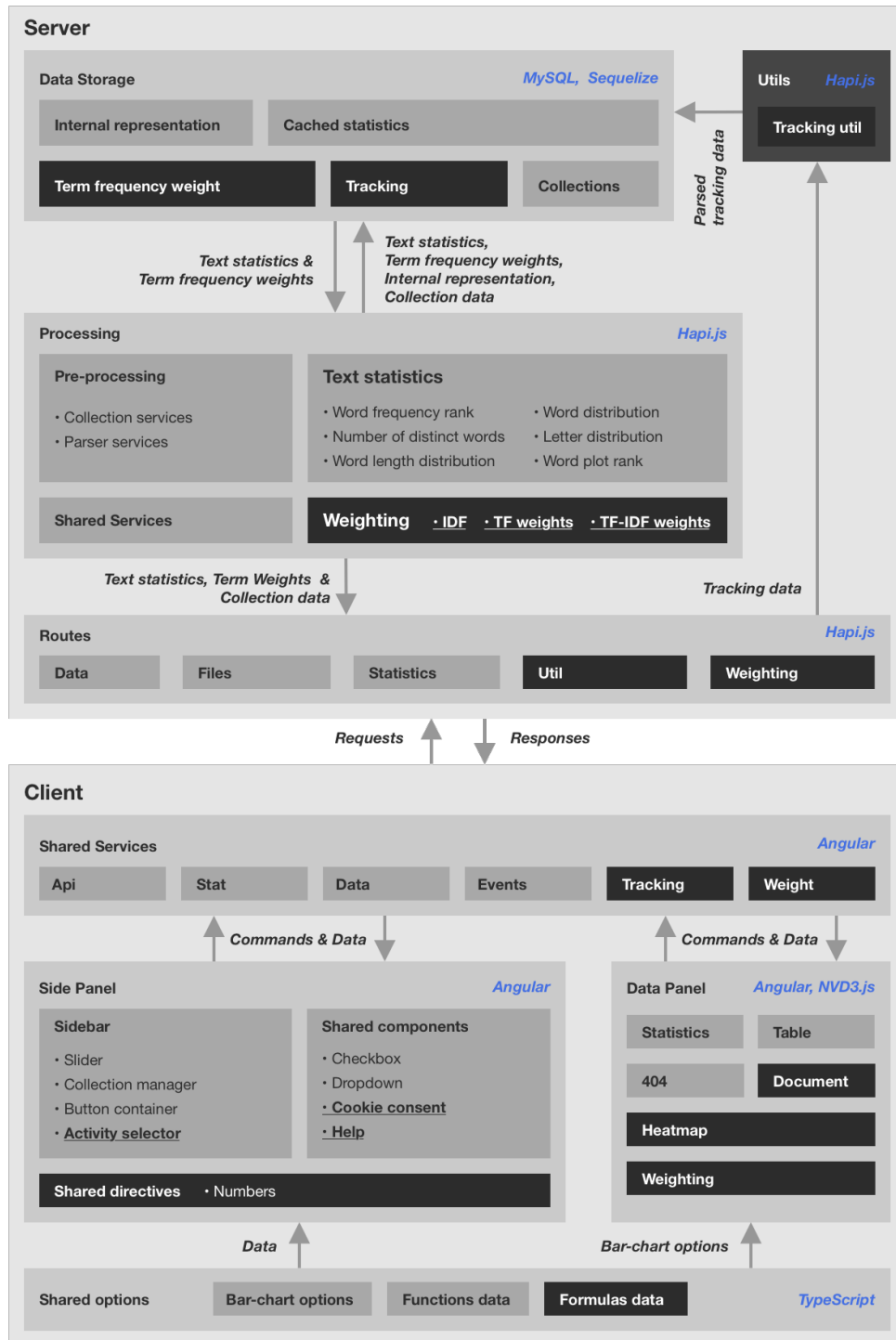


Figure 4.2.: The extended Infret architecture of the second version of Infret. Newly introduced components are displayed with black or dark grey blocks. The connections between components are depicted with lines where the arrows indicate the direction of data. The data transferred through them is described with italic text next to them.

4. Infret Prototype Version 2

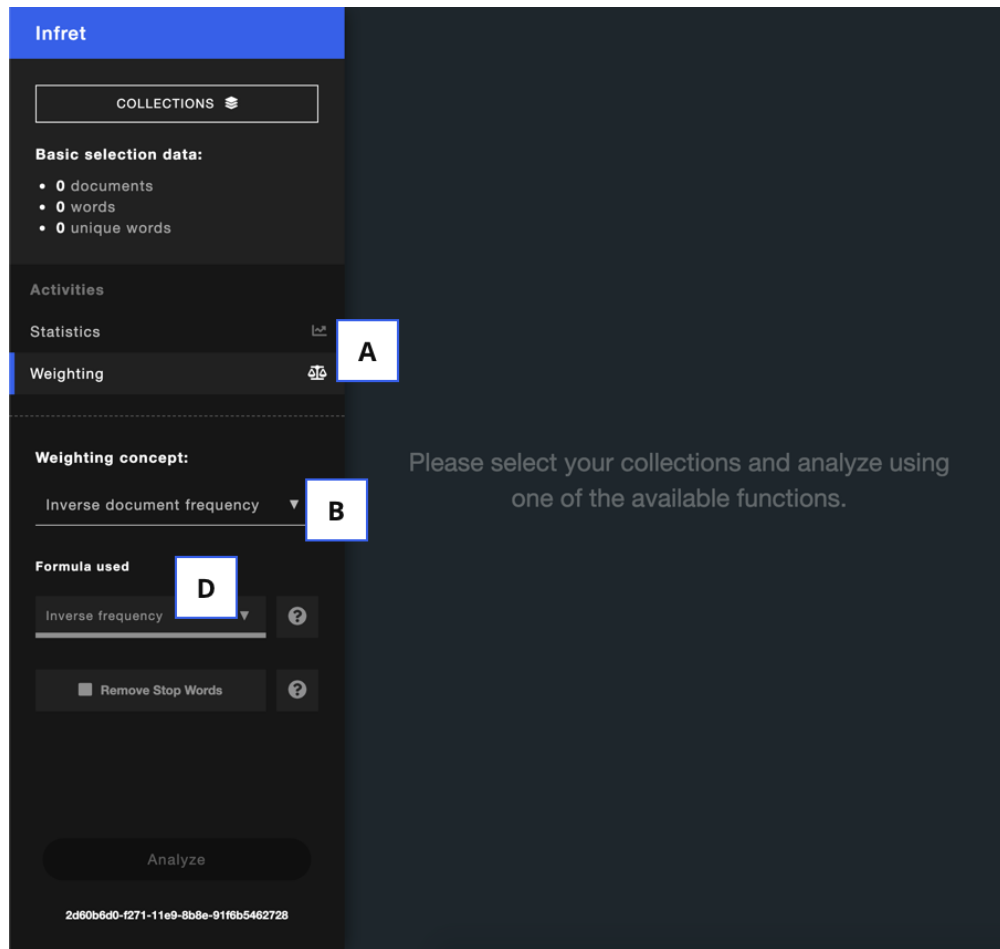


Figure 4.3.: Infret prototype version 2 UI overview

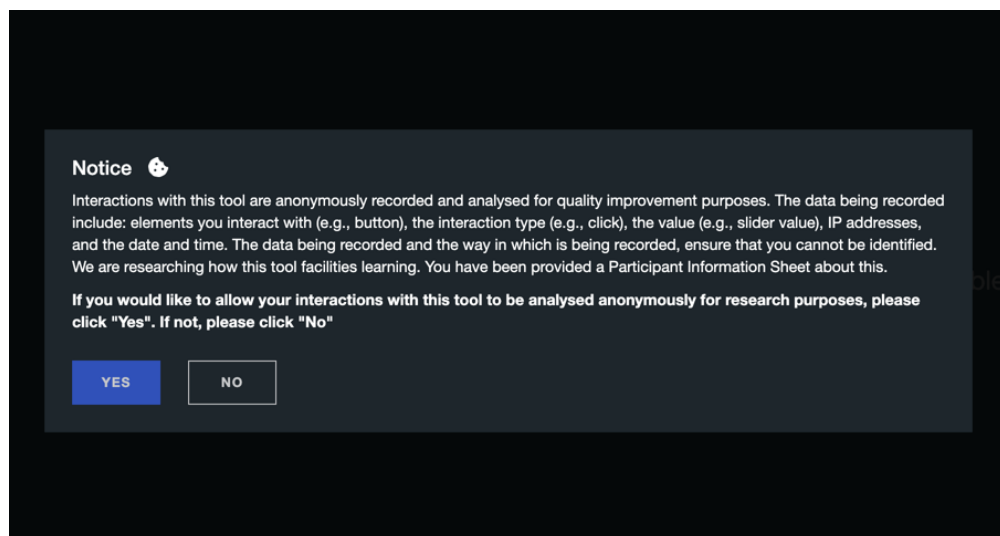


Figure 4.4.: Cookie notification

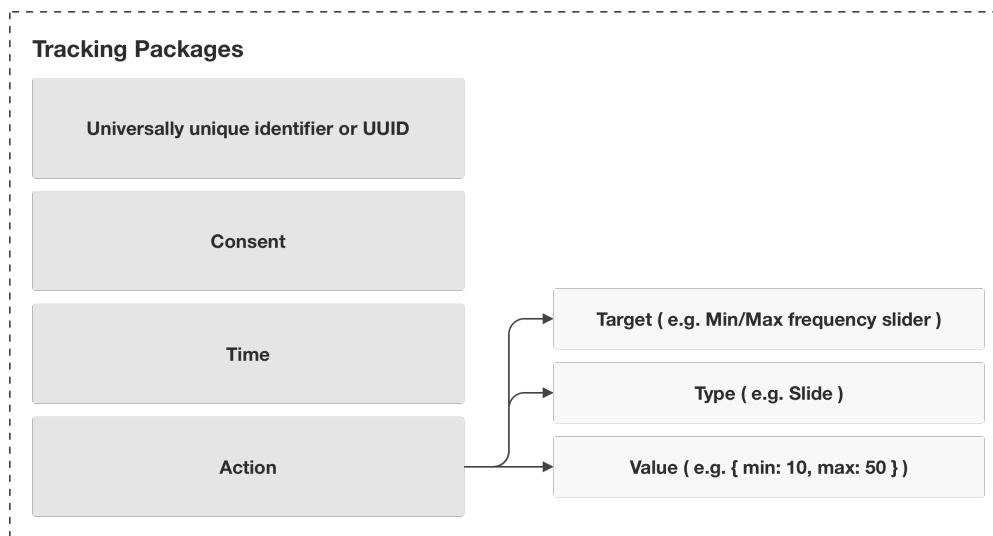


Figure 4.5.: Tracking data structure

Help modules The second version of Infret has additional features which help users understand various UI elements and concepts of Infret. Figure 4.6 displays a help tooltip being shown to a user whenever they hover over a UI element and explains the purpose of that element in the form of a short text. Furthermore, users can click buttons with a question mark to receive further explanation about the theory of a certain concept or further information about a certain feature (for example a list of stopwords) in form of a modal window. A modal with additional explanation for a concept seen in Figure 4.7 presents the user with the concept name, the formula of the concept (G) in case there is one, a short description of the purpose of the concept (H) and explanation for individual variables in the formula (I). Once the user is done refreshing their knowledge, they can close the modal by pressing the close button (F).

Term Frequency Weights To enable users the exploration of TF weighting in Infret, new UI elements which can be seen in Figure 4.8, had to be introduced. Once a user decides to investigate TF weights they can select one of the six TF weight formulae using a dropdown (J). They can additionally filter out words with a low or high term frequency using a slider element (K) and remove all stopwords (L). Once they trigger the analysis they are presented with a heatmap (M) of TF weights for all words and documents in the selected document collection. The columns represent unique words and rows represent unique documents. Due to performance issues, only 100 documents and 50 words can be shown in the heatmap at a time. Furthermore, due to the aforementioned limitations, a new pagination element (N) was introduced which enables users to navigate through all documents and words in form of pages where each document page contains 100 documents and each word page contains 50 words.

4. Infret Prototype Version 2

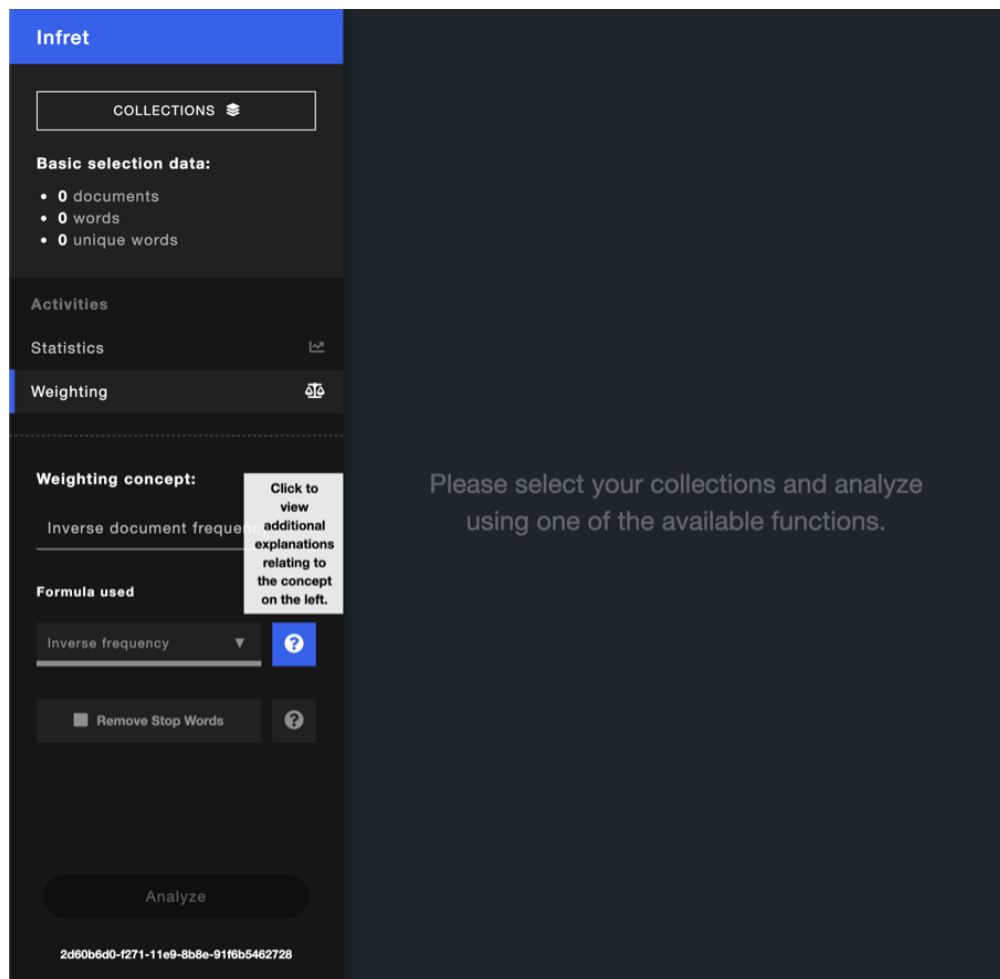


Figure 4.6.: Help tooltip shown when hovering over UI elements

Inverse frequency

CLOSE X

F

G $\log_2\left(\frac{N}{n_i}\right)$

H

The inverse document frequency is a function of the inverse of the number of documents in which a term occurs. It will be higher when the word occurs in few documents and lower when the word occurs in many documents. The selected formula is the basic inverse document frequency where the number of all documents is divided by the number of documents the word occurs in.

- N is the number of all documents in the collection.
- n_i is the number of documents in which the term appears.

I

Figure 4.7.: Help modal

4. Infret Prototype Version 2



Figure 4.8.: TF weight result example

The user can additionally investigate the content of each document by clicking on the document identifier number in the leftmost column. This brings up a modal seen in Figure 4.9 which displays the document identifier and word count (P) and the document content (R).

Inverse Document Frequency The next term weighting concept introduced in Infret is the IDF. There was no need to introduce new UI elements in the data panel for IDF, however, there was still a need to reuse new elements introduced in the side panel. An example of IDF results can be seen in Figure 4.10. Just like with TF weights users can select one of multiple formulae for IDF, filter words based on term frequency and remove stop words from the results.

Term Frequency - Inverse Document Frequency Weights Finally the TF-IDF weights, whose example results are displayed in Figure 4.11 are similar to TF weights. They use the same UI elements and require the same workflow as TF weights, however, the results are quite different as it can be seen by the heatmap in Figure 4.11 with only a few high weight values.

4.2. Development

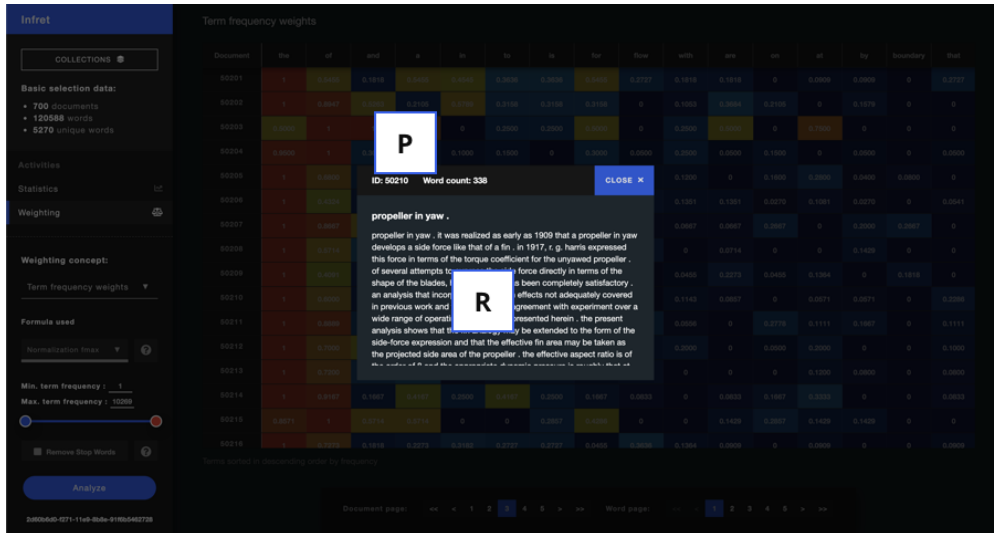


Figure 4.9.: Modal displaying content of a collection document

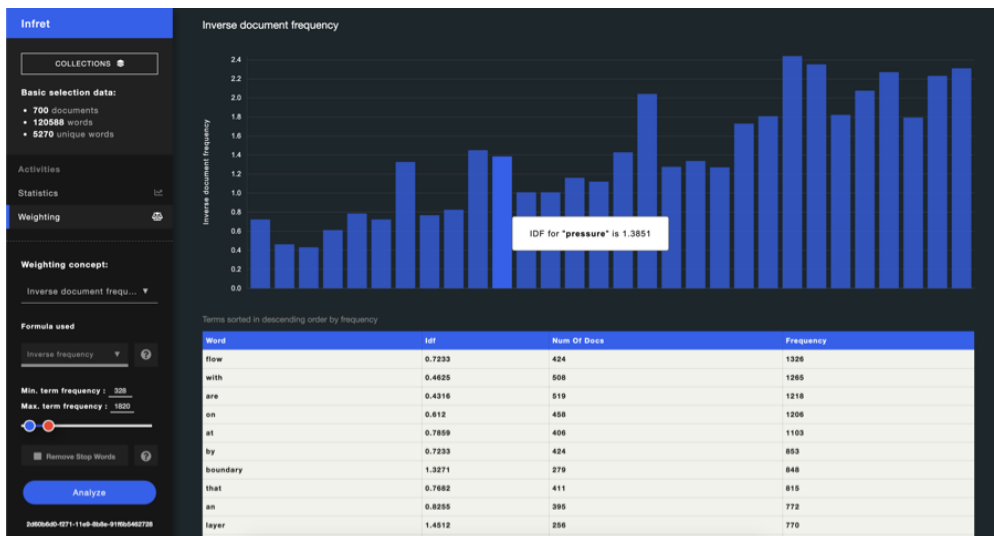


Figure 4.10.: Example IDF results

4. Infret Prototype Version 2

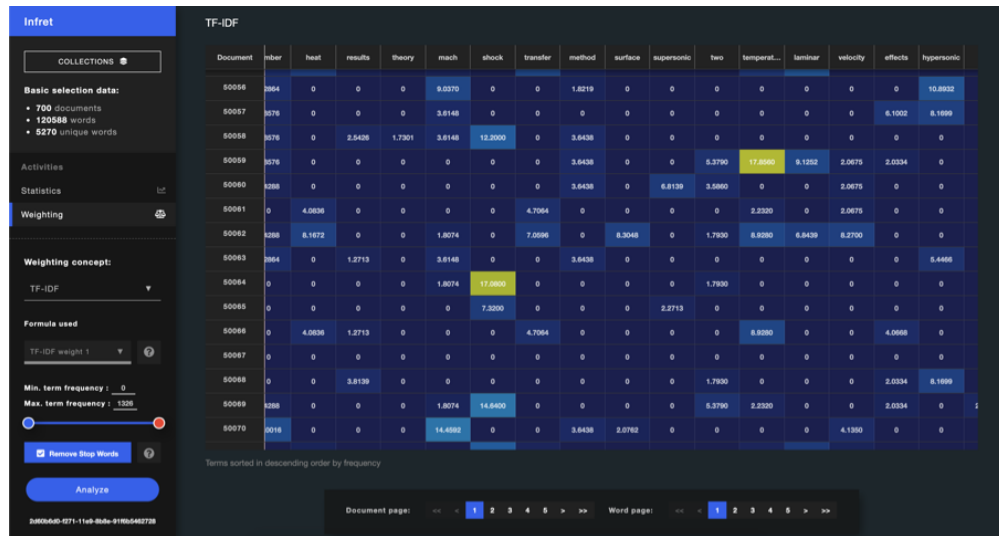


Figure 4.11.: Example TF-IDF weight results

4.3. Evaluation

Once Infret was extended with new features it was used by two student groups with different levels of prior knowledge as part of a class activity. Feedback data gathered during and after the activity was analysed to conclude the evaluation. This section describes the design of the study, the settings and instruments used to perform the study, the participant groups and finally, presents and discusses the findings. This section is based on and content has been partially taken from (Bobić, Cheong, et al., 2019)².

4.3.1. Study Design

The overall aim of the evaluation was to see how Infret helps students with different levels of prior knowledge understand IR concepts. Furthermore, we would like to see how the new UI changes and new features affect students experience while using Infret and how the evaluation results compare to the evaluation of the first prototype.

The second version of Infret was used as part of a class activity seen in Appendix C by two postgraduate student groups with different levels of prior knowledge to answer the aforementioned questions. The two groups received six tasks as part of their activity. Most tasks were similar, however, the experienced group with higher prior knowledge had to use their existing knowledge

²The original findings described in (Bobić, Cheong, et al., 2019) stated 22 experienced students completed the activities and filled out the survey. However, after another inspection of the data it became apparent there were only 20 who did both. That is why this chapter reports different numbers for the number of experienced students.

and advanced formulae to perform additional calculations as part of their tasks. Once the groups finished their activities they were provided with a link to an anonymous online survey seen in Appendix D. The survey consisted of demographic questions, prior knowledge questions, SUS (Brooke et al., 1996) questions, CES (Kay & Loverock, 2008) questions and general feedback questions. The results of this survey were later analysed using a variety of tools and approaches to answer the initial research questions.

4.3.2. Settings and Instruments

Both student groups received a short introduction on six activities seen in Appendix C which were most similar in the topic they were exploring, however, the activities for the experienced group required the use of formulae and skills they already learned while the activities for the novice group did not. An overview of the activities can be seen in Table 4.1.

Both groups firstly receive short instructions on how to access Infret and what document collection to select. Activity 1 instructs both groups to investigate and discuss letter distributions in their collection. Activity 2.1 requires the novice students to explore how the number of distinct words in their collection changes when the number of all words in the collection grows. The experienced students, on the other hand, have to investigate the word distributions of their collection and calculate various values connected to it as part of activity 2.2. Next both groups explore the distribution of selected words in single documents as part of activity 3. Activities 4 and 5 ask students to investigate the IDF values and TF / TF-IDF weights of words respectively. Finally, the student activities are concluded with activity 6 which is intended to help students refresh and review the knowledge about term weighting they gained in previous activities.

Table 4.1.: Structure of activities

Activity number	Activity focus	Target group
1	Letter distribution	Experienced & Novice
2.1	Number of distinct words	Novice
2.2	Word distribution in collection	Experienced
3	Word distribution in documents	Experienced & Novice
4	IDF	Experienced & Novice
5	TF / TF-IDF weights	Experienced & Novice
6	Term weighting reflection	Experienced & Novice

Overview of all activities performed by the two student groups. Taken and modified from (Bobić, Cheong, Filippou, Cheong, & Gütl, 2019).

In addition, to different activities, the two student groups also had different learning settings. The novice students were not graded and had lecturer

4. Infret Prototype Version 2

guidance while performing their activities. The lecturer shortly explained the theory behind each activity and how to navigate Infret to investigate the activity concept. The novice students had afterwards approximately 10 minutes to complete the activity. Once the time ran out the lecturer presented the findings and the goal of the activity. The students had 2 hours of exposure to Infret, during which they listened to lecturers explanations and completed the activities interchangeably. The main goal was to introduce novice students to new text statistic and term weighting concepts and enable them to explore these new concepts.

On the other hand, the experienced students were graded on their activity results, did not have any guidance, had 1 week to complete the activities in their private time and had to submit their answers. To help them navigate Infrets UI they were also provided with two annotated screenshots seen in Appendix E. Furthermore, experienced students were required to use the knowledge gained in the ISR course and perform calculations on data received while using Infret. The main goal of their activities was to enable them to get a deeper understanding of the learned concepts.

Once the activities were completed the two student groups received links to an anonymous online survey seen in Appendix D. The survey was identical for both groups and consisted of 6 question groups. The first and second groups were general demographic questions and prior knowledge evaluation questions respectively. To evaluate how Infret helps students understand various IR concepts the third group includes questions relating to individual concepts introduced in Infret. The fourth group is intended to measure the usability score of Infret and is comprised of 10 Likert-scale questions from SUS (Brooke et al., 1996). The fifth group is comprised of 12 questions from CES (Kay & Loverock, 2008) which are intended to explore students emotions while interacting with Infret. The survey concludes with a set of general feedback questions used for further improvement of Infret. An overview of all question groups can additionally be seen in Table 4.2.

4.3.3. Study participants

The study participants are comprised of two groups. The first group is a smaller group of experienced students, while, the second group is a larger group of novice students.

The 56 students from the novice group who completed the activities and filled out the survey were recruited in the first semester of 2019 from a postgraduate database design and development course at RMIT University in Melbourne. 21 (37.5%) of novice students were female and 35 (62.5%) were male. 41 (73.2%) of students were enrolled in the Master of Business Information Technology program. Furthermore, 35 (62.5%) novice students were between 18 and 24 years old, 20 (35.7%) were between 25 and 34 years old and 1 (1.78%) student

Table 4.2.: Structure of survey

Section	Section focus
1	Demographics questions
2	Prior knowledge questions
3	How Infret helps students understand specific areas of IR
4	System Usability Scale
5	Computer emotion Scale
6	General feedback questions

Overview of all survey question groups included in the survey filled out by both student groups once they completed the learning activities. Taken and modified from (Bobić, Cheong, Filippou, Cheong, & Gütl, 2019).

was between 35 and 44 years old. The novice students used Infret as part of their course at the start of the semester and did not have prior knowledge in IR. The 20 students from the experienced group who completed the activities and filled out the survey were recruited in the winter semester of the school year 2018/2019 from a postgraduate ISR course at the Graz University of Technology. 8 (40%) students were female and 12 (60%) were male. 11 (55%) experienced students were between 25 and 34 years old, 7 (35%) were between 18 and 24 years old and 2 (10%) were between 35 and 44 years old. The experienced students used Infret as part of their course at the end of their semester and had therefore already prior knowledge in IR concepts.

Students prior knowledge was measured using 4 different survey questions asking them about their prior knowledge in the following four topics:

- Text statistics
- Basic statistics concepts (probability, standard deviation, probability distribution, etc.)
- Basic data analysis approaches (observing data distribution, looking for correlation, etc.)
- Term weighting approaches

The students could rate their prior knowledge in the aforementioned topics using one of the following values:

1. Never heard of them or heard of them but don't know what they are.
2. I have an idea of what they are but don't know where or when to use them.
3. Can explain what they are but never used them.
4. Can explain what they are and used them before.

Statistics were calculated for the two student groups for each of the topics. A value of 1 (knowledge level 1) implies students have no prior knowledge about

4. Infret Prototype Version 2

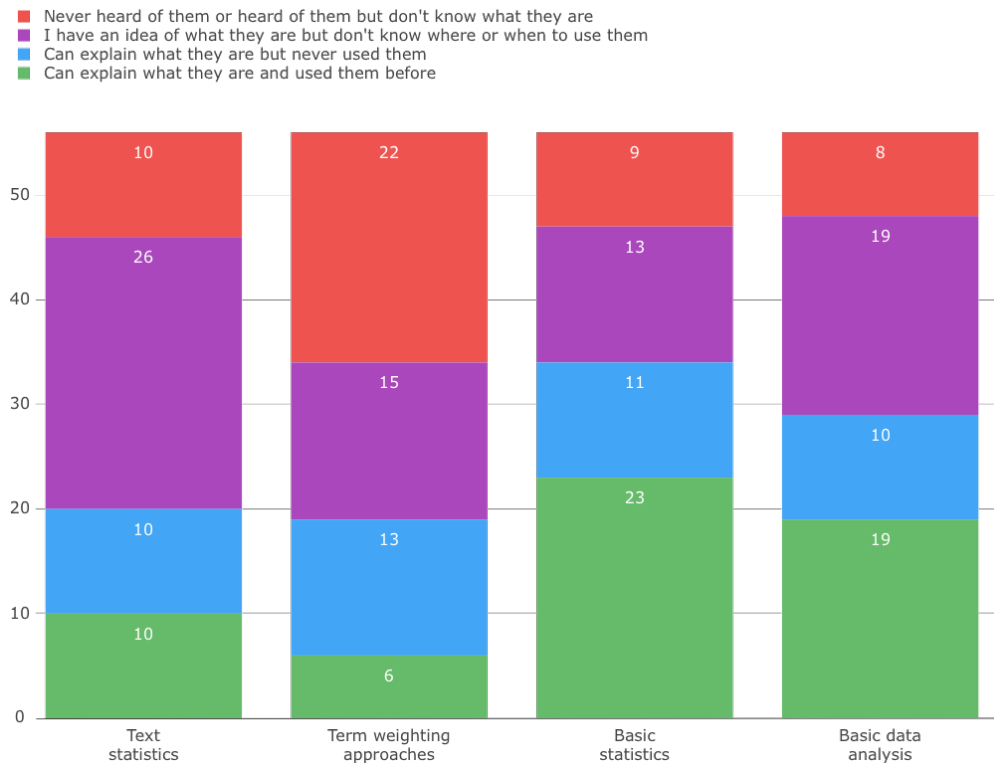


Figure 4.12.: Prior knowledge of novice students in 4 areas.

the topic or concept and a value of 4 (knowledge level 4) implies students have a lot of prior knowledge about the topic.

34 (60.7%) novice students could explain what basic statistics are (knowledge level 3 or 4) and 13 (23.2%) had an idea of what they are (knowledge level 2). When asked about basic data analysis approaches 29 (51.8%) had high prior knowledge (knowledge level 3 or 4) and 19 (33.9%) had low prior knowledge (knowledge level 2). 20 (35.7%) students could explain what text statistics are (knowledge level 3 or 4) and 26 (46.4%) had an idea of what they are (knowledge level 2). Finally, 19 (33.9%) students could explain what term weighting approaches are (knowledge level 3 or 4) and 15 (26.8%) had an idea of what they are (knowledge level 2). Detailed statistics about the answers of novice students can be seen in Figure 4.12.

On the other hand, 17 (85%) experienced students had high prior knowledge of basic statistics (knowledge level 3 or 4) and 3 (15%) had low prior knowledge in basic statistics (knowledge level 2). 15 (75%) of experienced students could explain what basic data analysis approaches are (knowledge level 3 or 4) and 5 (25%) had an idea of what they are (knowledge level 2). When it comes to text statistics 12 (60%) students could explain what they are (knowledge level 3 or 4) and 7 (35%) had an idea of what they are (knowledge level 2). Finally, when asked about term weighting approaches 11 (55%) students could explain what

4.3. Evaluation

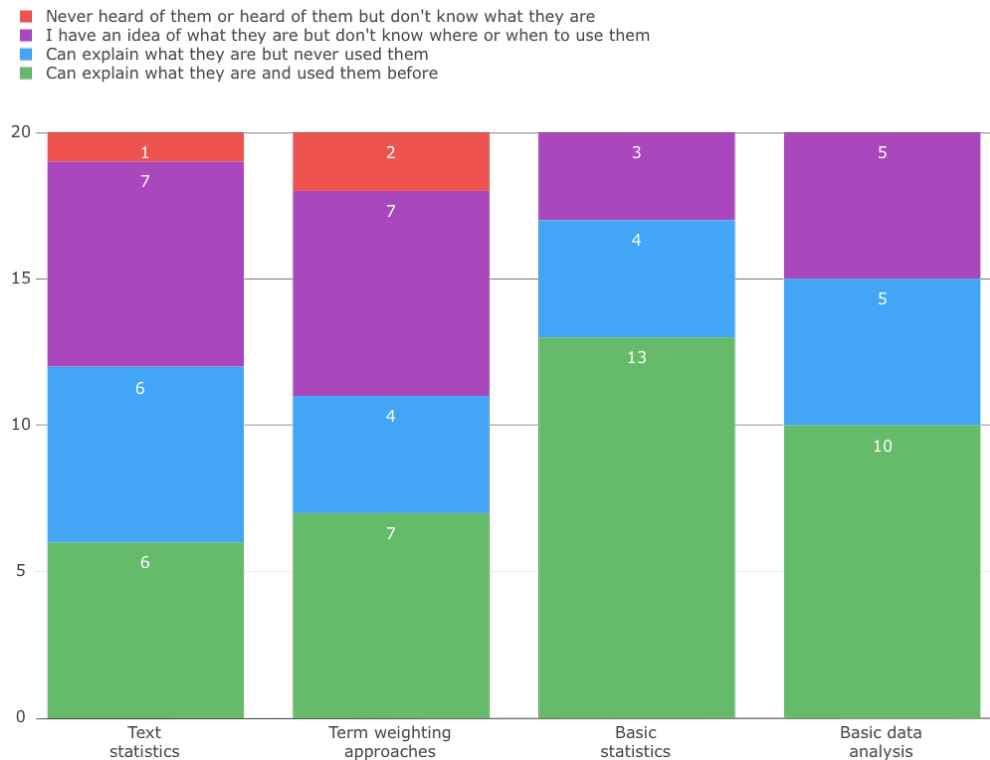


Figure 4.13.: Prior knowledge of experienced students in 4 areas.

they are (knowledge level 3 or 4) and 7 (35%) students had an idea of what they are (knowledge level 2). Detailed statistics about the answers of experienced students can be seen in Figure 4.13.

Finally, detailed statistics of prior knowledge values for both groups can be seen in Table 4.3. Although the values are close it can be concluded by comparing Figure 4.13 and Figure 4.12 that the novice group includes substantially more students with no prior knowledge of the aforementioned topics.

Table 4.3.: Prior knowledge statistic for experienced and novice students

Concept	Experienced	Novice
Text statistics	Mean: 2.85 SD: 0.93	Mean: 2.36 SD: 0.98
Term weighting approaches	Mean: 2.8 SD: 1.06	Mean: 2.05 SD: 1.03
Basic statistics	Mean: 3.5 SD: 0.76	Mean: 2.86 SD: 1.14
Basic data analysis	Mean: 3.25 SD: 0.85	Mean: 2.71 SD: 1.09

Overview of prior knowledge statistic for experienced and novice students. Taken and modified from (Bobić, Cheong, Filippou, Cheong, & Gütl, 2019).

4. Infret Prototype Version 2

4.3.4. Findings and Discussions

Hands-on Activities 19 (95%) experienced and 46 (82.14%) novice students agree they like experiments and hands-on activities. Furthermore, 18 (90%) experienced and 51 (91.07%) novice students claim experiments and hands-on activities help them better understand theory and methods. Almost no students disagreed with these two statements. Detailed answers can be seen in Table 4.4 and Table 4.5.

Table 4.4.: Experienced students opinion about experiments and hands-on activities

	Str. agree	Agree	Neutral	Disagree	Str. disagree
Q1	8 (40%)	11 (55%)	0 (0%)	1 (5%)	0 (0%)
Q2	10 (50%)	8 (40%)	1 (5%)	1 (5%)	0 (0%)

Experienced students answers to the following survey statements: Q1 - I like experiments and hands-on activities; Q2 - Generally, experiments and hands-on activities help me to better understand theory and methods.

Table 4.5.: Novice students opinion about experiments and hands-on activities

	Str. agree	Agree	Neutral	Disagree	Str. disagree
Q1	11 (19.64%)	35 (62.5%)	9 (16.07%)	0 (0%)	1 (1.79%)
Q2	10 (17.86%)	41 (73.21%)	5 (8.93%)	0 (0%)	0 (0%)

Novice students answers to the following survey statements: Q1 - I like experiments and hands-on activities; Q2 - Generally, experiments and hands-on activities help me to better understand theory and methods.

Additionally, most students (14 (70%) experienced and 39(69.6%) novice) had some experience with using interactive simulations to learn concepts before the activity with Infret. Finally, most novice students share the opinion that it is easier to grasp concepts by first having complete knowledge of the theory and then exploring the data and not the other way around. On the other hand, experienced students do not have a preferred order. It can be concluded that novice students should first attain theoretical background of concepts and then explore data with interactive e-learning tools.

Understanding IR Topics When asked if Infret helped students understand various IR topics most experienced students were divided between either agreeing that it helped them or expressing a neutral opinion. Their answers can be seen in Table 4.6. On the other hand, a strong majority of novice students expressed that Infret helped them better understand all of the IR topics that were part of the student activities. The novice students' answers are displayed in

Table 4.7. These findings indicate that Infret fulfills its goal of helping students better understand IR concepts in most cases.

Table 4.6.: How Infret helped experienced students understand various IR topics

Topic	Str. agree	Agree	Neutral	Disagree	Str. disagree
Text statistics	4 (20%)	7 (35%)	5 (25%)	2 (10%)	2 (10%)
TF weights	3 (15%)	7 (35%)	7 (35%)	1 (5%)	2 (10%)
IDF	3 (15%)	10 (50%)	4 (20%)	2 (10%)	1 (5%)
TF-IDF weights	5 (25%)	7 (35%)	6 (30%)	1 (5%)	1 (5%)

Detailed answers from experienced students to the questions with the following format “This tool has helped me to better understand and reflect aspects on text statistics / TF weights / IDF / TF-IDF weights”.

Table 4.7.: How Infret helped novice students understand various IR topics

Topic	Str. agree	Agree	Neutral	Disagree	Str. disagree
Text st.	7 (12.5%)	33 (58.93%)	12 (21.43%)	2 (3.57%)	1 (1.79%)
TF	10 (17.86%)	30 (53.57%)	11 (19.64%)	4 (7.14%)	1 (1.79%)
IDF	10 (17.86%)	31 (55.36%)	10 (17.86%)	3 (5.36%)	1 (1.79%)
TF-IDF	11 (19.64%)	26 (46.43%)	15 (26.79%)	2 (3.57%)	2 (3.57%)

Detailed answers from novice students to the questions with the following format “This tool has helped me to better understand and reflect aspects on text statistics / TF weights / IDF / TF-IDF weights”. It should be noted that one student did not answer the questions relating to text statistics and IDF. “Text st.” stands for “Text statistics” and “TF”, “TF-IDF” stand for “TF weights” and “TF-IDF weights” respectively.

Furthermore, when asked whether the various visualisations and the formula explanations helped them understand IR concepts better, most experienced and novice students agreed that line charts, bar charts and heat-maps helped them better understand concepts like text statistics, term weighting and IDF. However, when asked if the formula explanations helped them better understand the theory many (40%) experienced students answered with a neutral response. On the other hand, the majority (73.21%) of novice students agreed that formula explanations helped them more in understanding IR concepts than just exploring the data. These results further support the aforementioned conclusion that novice students need theoretical understanding before exploring data using interactive tools. Detailed answers of experienced students can be seen in Table 4.8 and of novice students in Table 4.9.

4. Infret Prototype Version 2

Table 4.8.: How visualisations and formula explanations helped experienced students understand various IR topics better

Question	Str. agree	Agree	Neutral	Disagree	Str. disagree
Q1	7 (35%)	5 (25%)	5 (25%)	1 (5%)	2 (10%)
Q2	7 (35%)	4 (20%)	6 (30%)	2 (10%)	1 (5%)
Q3	5 (25%)	6 (30%)	6 (30%)	2 (10%)	1 (5%)
Q4	7 (35%)	4 (20%)	5 (25%)	3 (15%)	1 (5%)
Q5	2 (10%)	6 (30%)	8 (40%)	1 (5%)	3 (15%)

Detailed answers from experienced students to questions regarding how visualisations and formula explanations helped them understand the IR topics better. Q1: "The bar chart visualisation helped me understand text statistics."; Q2: "The line chart visualisation helped me understand text statistics."; Q3: "The bar chart visualisation helped me understand idf."; Q4: "The heat-map helped me understand term weighting."; Q5: "The formula explanations when clicking the help button helped me understand the theory better than just exploring the data."

Table 4.9.: How visualisations and formula explanations helped novice students understand various IR topics better

Question	Str. agree	Agree	Neutral	Disagree	Str. disagree
Q1	8 (14.29%)	37 (66.07%)	8 (14.29%)	3 (5.36%)	0 (0%)
Q2	8 (14.29%)	35 (62.5%)	12 (21.43%)	1 (1.79%)	0 (0%)
Q3	14 (25%)	26 (46.43%)	13 (23.21%)	2 (3.57%)	1 (1.79%)
Q4	13 (23.21%)	30 (53.57%)	11 (19.64%)	1 (1.79%)	1 (1.79%)
Q5	13 (23.21%)	28 (50%)	11 (19.64%)	2 (3.57%)	1 (1.79%)

Detailed answers from novice students to questions regarding how visualisations and formula explanations helped them understand the IR topics better. Q1: "The bar chart visualisation helped me understand text statistics."; Q2: "The line chart visualisation helped me understand text statistics."; Q3: "The bar chart visualisation helped me understand idf."; Q4: "The heat-map helped me understand term weighting."; Q5: "The formula explanations when clicking the help button helped me understand the theory better than just exploring the data." It should be noted that one student did not answer Q5.

Usability Experienced students gave Infret a mean SUS score of 69.75 with a standard deviation of 11.67. Even though the received score is higher than the average which is 68 (Brooke, 2013) it is still lower than the mean SUS score of the first version of the Infret prototype which was 76.9 with a standard deviation of 11.8 (Bobić, Gütl, & Cheong, 2019). This drop in the usability score could be explained by the increased complexity of the UI introduced with new

features.

The novice student group gave Infret a lower mean SUS score of 57.68 with a standard deviation of 12. This score could indicate there is a larger problem with Infrets UI. However, it could also be attributed to the difference in the learning setting. The novice students had only 2 hours of exposure to Infret, were only shortly introduced to the concepts taught using Infret, were in a class environment and had lecturer guidance. These results might also support the aforementioned conclusions that novice students need a better understanding of the theory before exploring the data and need stronger support elements in the UI (such as formula explanations) to feel comfortable exploring the data. The detailed SUS data can be seen in Table 4.10. In conclusion, these results indicate that a larger focus should be placed on Infrets usability and Infrets support elements.

Table 4.10.: SUS Scores for the second version of the Infret prototype for experienced and novice students

	SUS Mean	SUS Standard deviation
Experienced	69.75	11.67
Novice	57.68	12

Emotions Just like the first version of Infret the survey for the second version consisted of 12 CES (Kay & Loverock, 2008) questions which were used to evaluate the students' emotions during their usage of Infret. The mean value and standard deviation of all measured emotions for experienced students can be seen in Table 4.11 and for novice students in Table 4.12. Furthermore, the values for experienced students have been visualised in Figure 4.14 and for novice students in Figure 4.15.

The results indicate that experienced users were more satisfied, less disheartened but experienced more frustration and angry than the students who were using the first version of Infret. Furthermore, the most dominant emotions were positive emotions such as satisfaction, excitement and curiosity.

On the other hand, the novice group experienced most emotions more often than the students who were using the first version of Infret. However, the difference in frequency compared to the students using the first version of Infret was higher for negative emotions than for the positive ones.

Finally when comparing experienced students to novice students it can be noticed they did not experience most negative emotions as often as the novice students. Especially when it comes to helplessness and nervousness which most experienced students didn't experience even once. This could be due to their prior knowledge while using Infret. Additionally, the experienced students weren't as excited as the novice students. These differences could again be

4. Infret Prototype Version 2

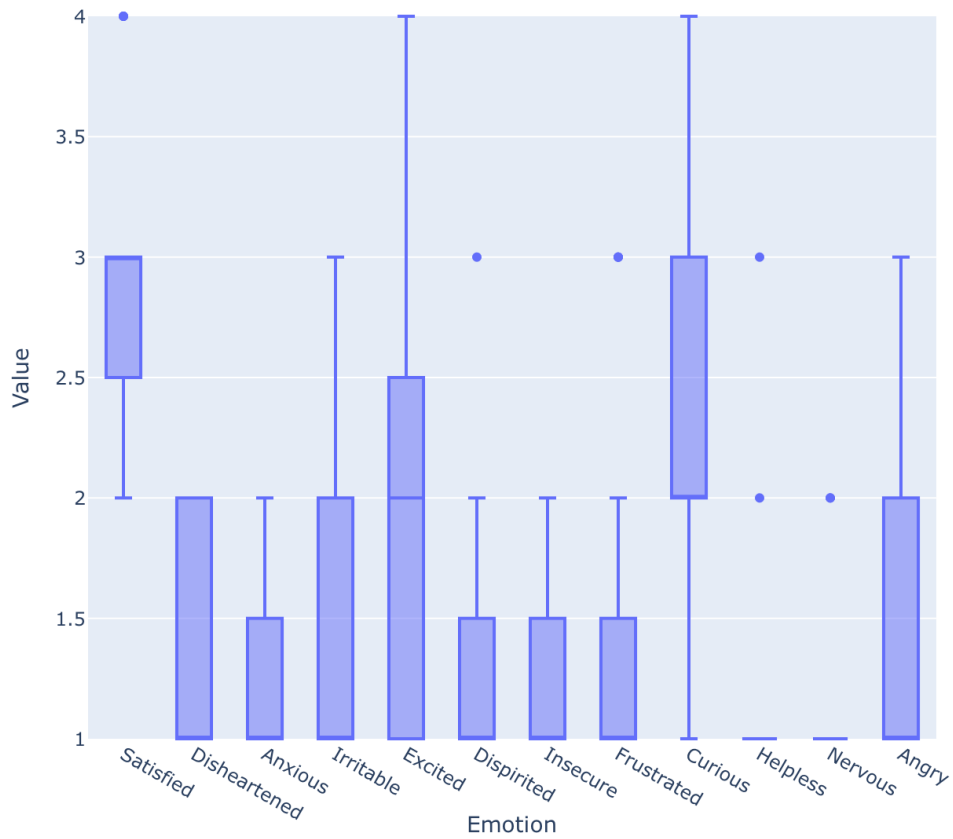


Figure 4.14.: All CES emotions of experienced students for the second version of the Infret prototype

due to the differences in the goals of the exercise for the two groups and the difference in their prior knowledge. The novice students were introduced to new topics while the experienced students already knew the concepts and were mostly repeating and using already learned knowledge.

4.3. Evaluation

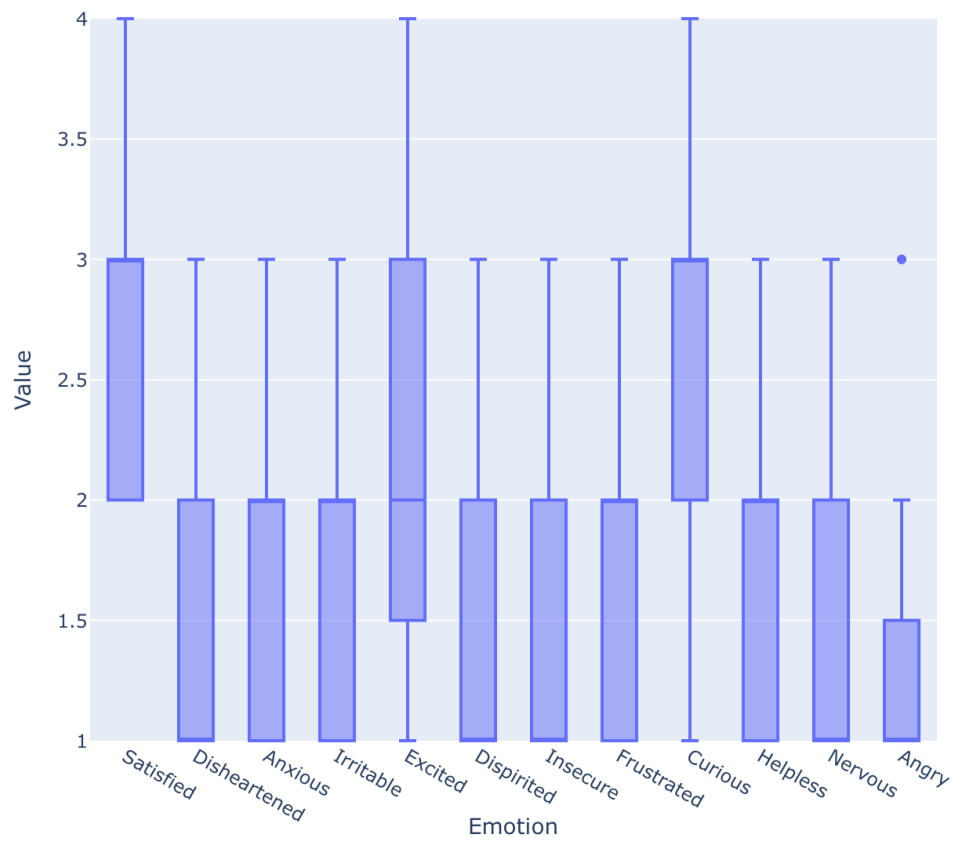


Figure 4.15.: All CES emotions of novice students for the second version of the Infret prototype

4. Infret Prototype Version 2

Table 4.11.: CES statistics for granular emotions of experienced students for the second version of the Infret prototype

Emotion	Mean	Standard deviation
Satisfied	2.9	0.64
Disheartened	1.3	0.47
Anxious	1.25	0.44
Irritable	1.35	0.59
Excited	2	0.86
Dispirited	1.3	0.57
Insecure	1.25	0.44
Frustrated	1.4	0.75
Curious	2.35	0.81
Helpless	1.15	0.49
Nervous	1.1	0.31
Angry	1.35	0.59

Table 4.12.: CES statistics for granular emotions of novice students for the second version of the Infret prototype

Emotion	Mean	Standard deviation
Satisfied	2.64	0.62
Disheartened	1.59	0.68
Anxious	1.73	0.67
Irritable	1.66	0.67
Excited	2.23	0.87
Dispirited	1.5	0.63
Insecure	1.57	0.68
Frustrated	1.7	0.71
Curious	2.68	0.86
Helpless	1.63	0.68
Nervous	1.5	0.63
Angry	1.29	0.53

The aforementioned observations can also be seen when observing the CES main emotions. The experienced students experienced more happiness but also more anger while the frequency of sadness and anxiety stayed the same as for the students using the first version of Infret. On the other hand, the novice students experienced all emotions more often than the students who were using the first version of Infret. However, they experienced anxiety much more often than the students who were using the first version of Infret. The

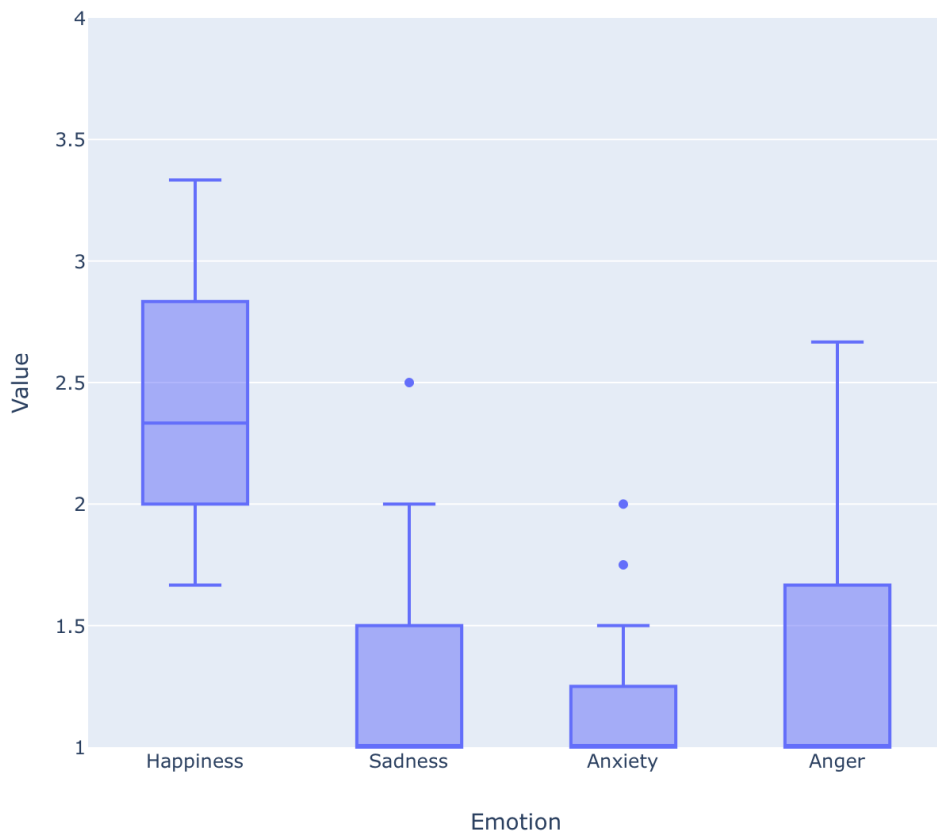


Figure 4.16.: CES four main emotions of experienced students for the second version of the Infret prototype

detailed values measured for all main emotions of experienced students can be seen in Table 4.13 and of novice students in Table 4.14. Additionally, the plotted values for experienced students are displayed in Figure 4.16 and for novice students in Figure 4.17

4. Infret Prototype Version 2

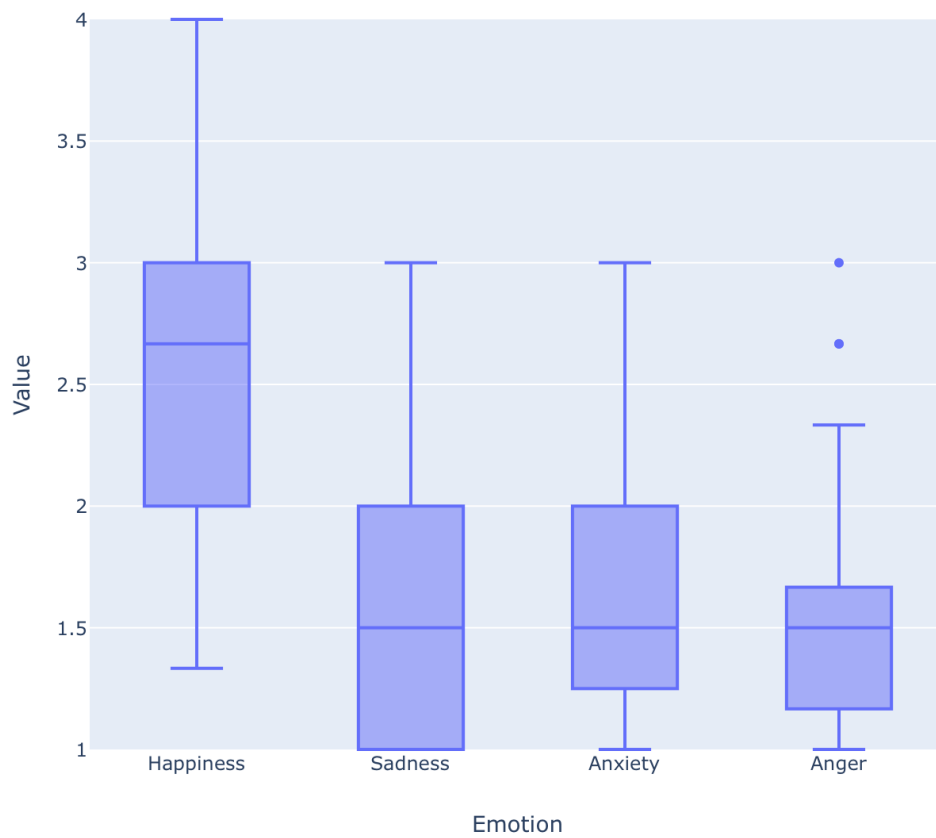


Figure 4.17.: CES four main emotions of novice students for the second version of the Infret prototype

Table 4.13.: CES statistics for four main emotions of experienced students for the second version of the Infret prototype

Emotion	Mean	Standard deviation
Happiness	2.42	0.53
Sadness	1.3	0.47
Anxiety	1.19	0.29
Anger	1.37	0.55

Table 4.14.: CES statistics for four main emotions of novice students for the second version of the Infret prototype

Emotion	Mean	Standard deviation
Happiness	2.52	0.62
Sadness	1.54	0.62
Anxiety	1.61	0.52
Anger	1.55	0.478

General feedback Even though the usability dropped and negative emotions increased 18 (90%) of experienced and 47 (83.93%) of novice students found Infret useful. Moreover, 17 (85%) experienced and 47 (83.93%) novice students agreed that Infret makes learning theory more interesting. The majority of both groups stated they would like to use such a tool for other subjects in IR. Furthermore, the results indicate that students think exploring the formulas using visualisations helped them more than just calculating the formulas on paper. Moreover, 8 (40%) experienced and 46 (82.14%) novice students agreed they are curious about other IR concepts that they haven't learned yet, while 7 (35%) experienced and 43 (76.78%) novice students would like to learn more about information retrieval. These results could indicate that Infret does spark curiosity for the majority of novice students who haven't been introduced to many IR concepts yet while it doesn't influence as many experienced students who are already familiar with IR theory and formed their opinion on the topic already in the past. Additionally, it can be concluded that Infret helps make the learning process more interesting for students with any prior knowledge level. Furthermore, when asked if coding the IR concepts would help them understand the theory the majority of both groups agreed. This could be another direction for Infret in the future which would further help students understand IR concepts. Statistics about answers to the aforementioned topics can be seen in Table 4.15 for experienced students and in Table 4.16 for novice students.

4. Infret Prototype Version 2

Table 4.15.: Answers to general feedback questions about Infret from experienced students

Question	Str. agree	Agree	Neutral	Disagree	Str. disagree
Q1	5 (25%)	13 (65%)	2 (10%)	0 (0%)	0 (0%)
Q2	7 (35%)	10 (50%)	3 (15%)	0 (0%)	0 (0%)
Q3	6 (30%)	7 (35%)	4 (20%)	3 (15%)	0 (0%)
Q4	5 (25%)	7 (35%)	6 (30%)	2 (10%)	0 (0%)
Q5	2 (10%)	6 (30%)	8 (40%)	3 (15%)	1 (5%)
Q6	1 (5%)	6 (30%)	10 (50%)	2 (10%)	1 (5%)
Q7	6 (30%)	7 (35%)	6 (30%)	1 (5%)	0 (0%)

Detailed answers from experienced students to general feedback questions about Infret. Q1: "I find the tool useful"; Q2: "This tool makes learning the theory more interesting"; Q3: "I would like to use such a tool for other subjects in ISR."; Q4: "Exploring the formulas using different visualisations helped me understand the theory more than just calculating them on paper."; Q5: "After using this tool I feel curious about other information retrieval concepts that I haven't learned yet."; Q6: "After using this tool I would like to learn more about information retrieval."; Q7: "Coding the IR concepts (term weighting, text statistics, ...) would have helped me to better understand the theory and methods."

Table 4.16.: Answers to general feedback questions about Infret from novice students

Question	Str. agree	Agree	Neutral	Disagree	Str. disagree
Q1	11 (19.64%)	36 (64.29%)	8 (14.29%)	1 (1.79%)	0 (0%)
Q2	10 (17.86%)	37 (66.07%)	7 (12.5%)	1 (1.79%)	1 (1.79%)
Q3	7 (12.5%)	37 (66.07%)	12 (21.43%)	0 (0%)	0 (0%)
Q4	8 (14.29%)	35 (62.5%)	13 (23.21%)	0 (0%)	0 (0%)
Q5	9 (16.07%)	37 (66.07%)	7 (12.5%)	2 (3.57%)	1 (1.79%)
Q6	9 (16.07%)	34 (60.71%)	11 (19.64%)	1 (1.79%)	1 (1.79%)
Q7	13 (23.21%)	34 (60.71%)	8 (14.29%)	1 (1.79%)	0 (0%)

Detailed answers from novice students to general feedback questions about Infret. Q1: "I find the tool useful"; Q2: "This tool makes learning the theory more interesting"; Q3: "I would like to use such a tool for other subjects in data and information retrieval."; Q4: "Exploring the formulas using different visualisations helped me understand the theory more than just calculating them on paper."; Q5: "After using this tool I feel curious about other information retrieval concepts that I haven't learned yet."; Q6: "After using this tool I would like to learn more about information retrieval."; Q7: "Coding the IR concepts (term weighting, text statistics, ...) would have helped me to better understand the theory and methods."

As a follow-up question students were asked in what other areas of ISR and for what other subjects they would like to use a tool like Infret. The participants mentioned they would like to use it for the following ISR areas and subjects:

- Search engines
- Boolean model
- Knowledge graphs
- Database design
- Statistics
- Data mining
- Anywhere where it is needed to understand new concepts
- Document ranking models
- Data analysis subjects

When asked what they liked about Infret students mentioned performance, visualisations, intuitive and appealing UI, simplicity, good user experience, that it helped them understand the theory better and that it motivated them to learn more about IR:

“Very intuitive. Apealing [sic] GUI.”

“I liked the visulization [sic] of the data a lot, charts, heat maps etc.”

“In general the tool helped me understand the concepts.”

“Generally nice user experience when using the tool and looking into data or performing tasks.”

“the usability, easy to use interface, different functionality”

“The given example of practical usage giving more motivation to work/learn in this field”

However, students also disliked areas of the tool such as the lack of data export and data filtering options, lack of search functionality, lack of copy functionality for some of the data, heatmap page navigation:

“I think there needs to be a method to filter the words which we need to find. Scrolling through pages is too time consuming.”

“Not being able to copy and paste”

“Results cannot be processed well: No export, no easy handling to copy data, no filtering, no search”

4. Infret Prototype Version 2

“There aren’t enough possibilities to filter the data. It is not always possible to copy the content into excel.”

“the width / style of the text, reading got exhausting with longer text passages”

Moreover, possibilities for improvements such as the addition of filtering function, copy function, fixing the term frequency filter, addition of a search function, improving UI for a better overview of possible functions, the addition of data highlighting option and more were identified by the students:

“highlight a particular data or combination of data”

“export data to other softwares.”

“add search button to find words easily”

“Provide filter that updates results, follow web best practices for scrolling, data presentation”

“There aren’t enough possibilities to filter the data. It is not always possible to copy the content into excel.”

“If you select a range for min/max term frequency before you clicked on “Analyse”, it doesn’t get applied. Fix the bug.”

“Better structuring of the input and make it more intuitive. I did not realize that I would have had the chance to view the formulars.”

“loading of charts and heatmaps for specific words - easier to see needed results without endless scrolling and searching among other terms”

The aforementioned results indicate students were mostly happy with the general usability of Infret and had mostly issues with the lack of features. This indicates that the SUS scores need to be interpreted carefully and not be taken as the absolute truth.

To sum up, both groups of students enjoyed using Infret and were able to gain new knowledge of IR concepts by using it. It completed its main objective of helping students better understand the concepts of text statistics and term weighting. The newly complex UI received a lower usability score and students experienced negative emotions more often than with the first version of Infret. Novice students experienced all emotions more often and indicated the addition of assistance elements such as the formula explanations helped them by understanding concepts and navigating Infret better. These results in addition to the identified possible improvements such as the addition of search functionality, filtering and exporting of data will be taken into account to build the next version of Infret.

4.4. Summary

The second version of Infret was conceptualised, developed and evaluated based on the findings and evaluation results of the first version of Infret (Bobić, Gütl, & Cheong, 2019). The second version introduced UI improvements, features such as TF weights, IDF, TF-IDF weights and with them new visualisations and the possibility to track user interactions. A novice and an experienced group of students used Infret as part of a learning activity. The novice group was a group of students who were at the start of their semester in a database design and development course and did mostly not have prior knowledge of IR. On the other hand, the experienced group was at the end of their semester in an ISR class where they familiarised themselves with many IR concepts. The six learning activities guided students to explore various IR concepts while interacting with Infret. Once they completed the activities, they filled out an anonymous survey for the evaluation of their experience with Infret. The results indicate the second version was received well. However, the usability scores were lower than for the first version. On the other hand, based on the feedback received from students Infrets user experience was well received and the most common complaint was the lack of features. Additionally, students experienced negative emotions more often than with the first version. However, positive emotions were still prevalent. Based on the overall feedback, multiple potential improvements were identified which might improve students experience further and will be implemented in the next version of Infret. Additionally, as part of future work, the interaction data gathered from students interaction with Infret will be analysed in detail to get a better understanding of how students interact with Infret.

5. Lessons Learned

Unlike the previous chapters which were describing stages of my research work, this chapter will provide an insight into my views and lessons learned while doing this work.

When starting my initial literature research I noticed it was hard finding relevant papers for my work. However, the more time I put into it the more it became apparent that there are plenty of papers and they are well connected and built on the same base layers of scientific work. Later when searching for existing solutions I noticed a similar pattern where the more literature I read the easier it became to find new relevant literature. While reading research papers I learned to appreciate the strict structure of scientific work which made it easier to initially skim through the text and conclude if the paper is relevant for my work. This motivated me to also hold myself to the same standards and try to make my written work as structured as possible. Additionally, I gained a new appreciation for the pedagogic process and for the years of theory and research that went into it and helped create new approaches such as TEAL and MAL. It made me realise how structured the pedagogic process is and how much time a lecturer has to invest to attempt to address the various difficulties students face while studying.

When it came to the development process I was at first not sure how the conceptual architecture and the requirements would help me while developing. Initially, I thought it would be easier to start immediately working on the prototype since it was a small application. However, it became apparent quickly that the guidelines set by the requirements and the conceptual architecture help immensely while planning and also prioritising various features of any app. Furthermore, it proved to be much harder to pick the appropriate new features for the second version of Infret. Some of the IR concepts would not benefit from being visualised or provided interactively while others would be too time-consuming for implementation and would not add enough additional value for the initial iterations of Infret. While developing I stumbled on many issues with code modularity and trying to keep the code base readable. Although Angular offers a code style guide it is still easy to make a wrong turn while developing. Moreover, using the git workflow with multiple feature branches and a develop branch proved to be useful as Infret grew larger despite my initial doubts. It enabled me to look up a previous version of Infret, revert features, track my progress and keep an overview of Infrets features.

Finally having a set of defined research questions together with the main

5. Lessons Learned

research question proved to be very useful while structuring the study. The research questions enabled me to organise the study with ease and focus on key aspects which would help me answer my research questions while creating the survey. During the creation of activities, I gained additional insight and appreciation for pedagogic work and realised it is hard to create new activities with solutions even if you are knowledgeable in your field. Even though the survey questions were carefully structured it was apparent from the lack of answers that some of the open-ended questions such as specific questions asking how each of the visualisations helped students understand the concepts could be left out or just grouped into one question. This would have help students complete the survey faster and at the same time make the process of analysing the answers shorted. Therefore I will take greater care when creating such surveys in the future.

In conclusion I realised that carrying out a research project requires a variety of skills, a lot of work and a lot of dedication. At the start of this project, I saw software development as the main part of the project. However, in the end, I knew it is only a small part of the puzzle. Research is a complex and time-demanding process, however, it is also very rewarding and insightful.

6. Conclusion and Future Work

This thesis describes the conceptualisation, development and evaluation of the prototype and of the second version of a Web-based IR tool called Infret. Infret is intended to be used as a visual component of a MAL-based (Pirker et al., 2014) class which would help students understand various IR topics.

A literature survey was conducted and provided an overview of existing pedagogic approaches. Selected interactive approaches built on ideas of constructivism and exploration were shown to help increase students knowledge retention and focus during lectures. Based on these approaches a new approach called motivational active learning (MAL) (Pirker et al., 2014) was developed to introduce interactivity and exploration to students in an ISR class with the use of gamification components. However, the MAL approach was missing an interactive visual component. To gain an insight into the current state of interactive visual and exploratory tools a survey of selected existing tools has been conducted. It was concluded that the majority of tools focus either on a reimplementing of IR concepts by offering a code framework, focus on query performance analysis, are too complex for in-class usage, are not focused on visualisation and interactivity or do not represent the basic concepts of IR. Based on the aforementioned findings it was concluded that a new tool for IR should be built.

The first Infret prototype was conceptualised based on the findings of the aforementioned literature survey and the evaluation findings of the ISRApp (Ziessler, 2018). The main aim of the prototype was to help students explore and learn various text statistics concepts using one of the selected text collections while serving as an interactive visualisation and exploration component of MAL. A set of functional and non-functional requirements was defined. Based on these requirements a conceptual architecture was designed to guide the development process. Infret's final architecture resembles the conceptual architecture and provides students with a simple UI for exploring various text statistics concepts by administering them to one of the available document collections. Infret was used in an ISR class as part of a student activity by a class of 25 masters students. Once they completed the activities they were provided with a survey whose results were later analysed. The evaluation results of the first version indicate that the Infret prototype has good usability and that students experience mostly positive emotions while using it. Additionally, most students would like to use it for other areas of ISR. Even though Infret was well received multiple potential improvements were identified. The process and results of the evaluation were

6. Conclusion and Future Work

published as part of (Bobić, Gütl, & Cheong, 2019).

The aforementioned findings were used to create the second version of Infret whose aim is to help students explore and understand term weighting concepts such as TF weights, IDF and TF-IDF weights in addition to text statistics. First, a set of functional requirements was defined. The conceptual architecture of the first version of Infret was extended to accommodate the new requirements. The second version of Infret was extended with the new features and its architecture continues to reflect the extended conceptual architecture. Furthermore, other minor improvements were made to the codebase. The second iteration of Infret was used and evaluated by a novice group of students with lower prior knowledge in IR and an experienced group of students with higher prior knowledge in IR. The results indicate Infret successfully helped both groups learn new IR concepts or improve their already existing knowledge of IR concepts. Even though the usability score was lower for both groups than the score of the Infret prototype and both groups experienced more negative emotions than with the previous version they both enjoyed Infrets user experience and agreed it helped them with the understanding of IR concepts. Furthermore, many potential improvements have been identified and will be used for future versions of Infret. The details of this process and evaluation were published as part of (Bobić, Cheong, et al., 2019)

Future versions of Infret will be expanded with additional features such as data export options, keyword filtering and search functionality, onboarding guide through the UI for new users, ability to code custom implementations of concepts such as text statistics within Infret and more. Additionally, a wider range of IR concepts will be introduced to Infret to help students learn and explore advanced topics which would benefit from an exploratory visual presentation. Furthermore, a new administration panel could be introduced which would provide lecturers with an overview of the currently active users and their progress through learning activities. Moreover to define automatically graded activities and lock certain areas of Infret for these activities Infret could be connected to an external LMS such as Moodle. Finally, Infret could be expanded with a variety of gamification features such as leaderboards for comparing custom implemented functions and guided learning of new concepts in Infret where students would collect achievements for each learned concept. Just like the existing version of Infret, new iterations will be evaluated in order to monitor Infrets progress and investigate how it can help students learn a wider range of IR concepts in a more efficient manner.

Bibliography

- Alberts, P. P., Murray, L. A., Griffin, D. K., & Stephenson, J. E. (2007). Blended learning: Beyond web page design for the delivery of content. *Blended Learning*, 53.
- Baeza-Yates, R., Ribeiro-Neto, B. et al. (1999). *Modern information retrieval*. ACM press New York.
- Bates, A. T. (2005). *Technology, e-learning and distance education*. Routledge.
- Beichner, R. J. & Saul, J. M. (2003). Introduction to the scale-up (student-centered activities for large enrollment undergraduate programs) project. *Proceedings of the International School of Physics "Enrico Fermi," Varenna, Italy*. <http://www.ncsu.edu/per/scaleup.html> (accessed 7 June 2005).
- Białecki, A., Muir, R., Ingersoll, G., & Imagination, L. (2012). Apache lucene 4. In *Sigir 2012 workshop on open source information retrieval* (p. 17).
- Bobić, A., Cheong, C., Filippou, J., Cheong, F., & Gütl, C. (2019). Infret: Enhancing a tool for explorative learning of information retrieval concepts. In *Icl2019 proceedings*. Advances in Intelligent Systems and Computing. Presented at ICL conference 2019 and to be published in Springer Proceedings.
- Bobić, A., Gütl, C., & Cheong, C. (2019). Infret: Preliminary findings of a tool for explorative learning of information retrieval concepts. In *Icbl2019 proceedings*. Lecture Notes in Networks and Systems. Presented at ICBL conference 2019 and to be published in Springer Proceedings.
- Bonwell, C. C. & Eison, J. A. (1991). *Active learning: Creating excitement in the classroom*. 1991 ashe-eric higher education reports. ERIC.
- Brooke, J. et al. (1996). Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194), 4-7.
- Brooke, J. (2013). Sus: A retrospective. *Journal of usability studies*, 8(2), 29-40.
- Bunce, D. M., Flens, E. A., & Neiles, K. Y. (2010). How long can students pay attention in class? a study of student attention decline using clickers. *Journal of Chemical Education*, 87(12), 1438-1443.
- Calado, P., Cardoso-Cachopo, A., & Oliveira, A. L. (2007). Ir-base: An integrated framework for the research and teaching of information retrieval technologies. In *Int. workshop on teaching and learning of inf. retrieval*.
- Cummings, K., Marx, J., Thornton, R., & Kuhl, D. (1999). Evaluating innovation in studio physics. *American journal of physics*, 67(S1), S38-S44.

Bibliography

- De Freitas, S. & Neumann, T. (2009). The use of 'exploratory learning' for supporting immersive learning in virtual environments. *Computers & Education, 52*(2), 343–352.
- Dori, Y. J. & Belcher, J. (2005). How does technology-enabled active learning affect undergraduate students' understanding of electromagnetism concepts? *The journal of the learning sciences, 14*(2), 243–279.
- Dori, Y. J., Hult, E., Breslow, L., & Belcher, J. W. (2007). How much have they retained? making unseen concepts seen in a freshman electromagnetism course at mit. *Journal of Science Education and Technology, 16*(4), 299–323.
- Efthimiadis, E. N. & Freier, N. G. (2007). Ir-toolbox: An experiential learning tool for teaching ir. In *Proceedings of the 30th annual international acm sigir conference on research and development in information retrieval* (pp. 914–914). ACM.
- Fang, H., Wu, H., Yang, P., & Zhai, C. (2014). Virlab: A web-based virtual lab for learning and studying information retrieval models. In *Proceedings of the 37th international acm sigir conference on research & development in information retrieval* (pp. 1249–1250). ACM.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, 111*(23), 8410–8415.
- Garrison, D. R. & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The internet and higher education, 7*(2), 95–105.
- Güzer, B. & Caner, H. (2014). The past, present and future of blended learning: An in depth analysis of literature. *Procedia-social and behavioral sciences, 116*, 4596–4603.
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American journal of Physics, 66*(1), 64–74.
- Halttunen, K. & Sormunen, E. (2000). Learning information retrieval through an educational game. is gaming sufficient for learning? *Education for Information, 18*(4), 289–311.
- Henrich, A. & Sieber, S. (2009). Blended learning and pure e-learning concepts for information retrieval: Experiences and future directions. *Information Retrieval, 12*(2), 117–147.
- Johnson, D., T. Johnson, R., & Smith, K. (1998). Active learning: Cooperation in the college classroom. *The Annual Report of Educational Psychology in Japan, 47*. doi:10.5926/arepj1962.47.0_29
- Johnson, R. T., Johnson, D. W., & Stanne, M. B. (1986). Comparison of computer-assisted cooperative, competitive, and individualistic learning. *American educational research journal, 23*(3), 382–392.

- Kay, R. H. & Loverock, S. (2008). Assessing emotions related to learning new software: The computer emotion scale. *Computers in Human Behavior*, 24(4), 1605–1623.
- Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. FT press.
- López-García, R. & CACHEDA, F. (2011). A technical approach to information retrieval pedagogy. In *Teaching and learning in information retrieval* (pp. 89–105). Springer.
- Lopez-Herrera, A., Alonso, S., Cabrerizo, F., Porcel, C., Cobo, M., & Herrera-Viedma, E. (2010). Using a visual tool to guide students in the complex process of learning fuzzy weighted information retrieval systems. In *Inted2010 proceedings* (pp. 361–370). IATED.
- Manning, C., Raghavan, P., & Schütze, H. (2010). Introduction to information retrieval. *Natural Language Engineering*, 16(1).
- Ounis, I., Amati, G., Plachouras, V., He, B., Macdonald, C., & Johnson, D. (2005). Terrier information retrieval platform. In *European conference on information retrieval* (pp. 517–519). Springer.
- Pirker, J., Riffnaller-Schiefer, M., & Gütl, C. (2014). Motivational active learning: Engaging university students in computer science education. In *Proceedings of the 2014 conference on innovation & technology in computer science education* (pp. 297–302). ACM.
- Prince, M. (2004). Does active learning work? a review of the research. *Journal of engineering education*, 93(3), 223–231.
- Richardson, V. (2003). Constructivist pedagogy. *Teachers college record*, 105(9), 1623–1640.
- Rowe, M. B. (1983). Getting chemistry off the killer course list. ACS Publications.
- Ruhl, K. L., Hughes, C. A., & Schloss, P. J. (1987). Using the pause procedure to enhance lecture recall. *Teacher education and special education*, 10(1), 14–18.
- Salton, G. & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5), 513–523. doi:10.1016/0306-4573(88)90021-0
- Wilhelm-Stein, T., Kahl, S., & Eibl, M. (2017). Teaching the information retrieval process using a web-based environment and game mechanics. In *Proceedings of the 40th international acm sigir conference on research and development in information retrieval* (pp. 1293–1296). ACM.
- Ziessler, F. (2018). *A framework for e-learning in the field of information retrieval* (Master's thesis, Graz University of Technology).

Appendix

Appendix A.

Infret Prototype Version 1 - Activities

Experimentation on text statistics

The goal of this practical experimentation is to investigate text statistics from a pre-existing document collection and compare the results with the theory covered in lecture 5. For this exercise a tool is provided; unpack and run the tool (see also readme file). Select in the tool the document collection: Use CranfieldCollection (1400 documents) for even student ID number and CranfieldCollectionHalf (700 documents) for odd student ID number. Administer the following tasks and document your findings in a report.

a. Overall statistics

Give an overview about the document collection.

b. Single letter distribution

Administer the probability distribution on the level of single characters. Provide your findings in a table. Discuss the findings and find 2 application scenarios for IR making use of the letter distribution.

c. Word distribution

Investigate the word distribution of the first 40 words (rank, word count, percentage of total). Document the findings in a table. Show in a graph how “ $r * p_r$ ” varies.

Apply

$$f_r = f(r) = f_1 r^{-\alpha}$$

and find for the factor “alpha” the best value (least square sum error).

Apply

$$p_r = \frac{f_r}{f} = \frac{A}{r}$$

Appendix A. Infret Prototype Version 1 - Activities

and discuss the results. Find the best value for “A” (least square sum error).

d. Number of words occurring n-times

Apply

$$\frac{R_i}{V} = \frac{1}{f_i(f_i + 1)}$$

and compare the values with the practical findings in a table for N= 1 ... 15.

e. Number of distinct words

Apply

$$V = Kf^\beta$$

and estimate K and “beta” for the given results.

Appendix B.

Infret Prototype Version 1 - Survey

Thank you for participating in this survey. This survey is being conducted to optimize the experimentation tool and the learning model for further courses. The survey results are anonymised and will be solely used for evaluation and research purposes. Your participation is completely voluntary. Completion of this survey implies your consent for the purposes stated above.

Gender

- Female
- Male
- No answer

Age

Number entry field.

Field of Study

- Computer Science
- Software Development and Business Management
- Information and Computer Engineering
- Mathematics
- Doctoral programme
- Other

Rate the following statements between “I strongly disagree (1)” and “I strongly agree (5)”

- I like experiments and hands-on activities.
- Generally, experiments and hands-on activities help me to better understand theory and methods.
- This tool has helped me to better understand and reflect aspects on text statistics.

Appendix B. Infret Prototype Version 1 - Survey

- Coding the text statistics function would have helped me to better understand the theory and methods.
- I would like to use such a tool for other subjects in ISR.

For which other subjects in ISR would you like to use it?

Free form entry field.

For which other subjects could such an approach be used?

Free form entry field.

What I liked most:

Free form entry field.

What I did not like at all:

Free form entry field.

What could be improved:

Free form entry field.

Rate the following statements between “I strongly disagree (1)” and “I strongly agree (5)”

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

Rate the following statements between “None of the time (1)” and “All of the time (4)”

- Using this software, I feel ... Satisfied

- Using this software, I feel ... Disheartened
- Using this software, I feel ... Anxious
- Using this software, I feel ... Irritable
- Using this software, I feel ... Excited
- Using this software, I feel ... Dispirited
- Using this software, I feel ... Insecure
- Using this software, I feel ... Frustrated
- Using this software, I feel ... Curious
- Using this software, I feel ... Helpless
- Using this software, I feel ... Nervous
- Using this software, I feel ... Angry

Appendix C.

Infret Prototype Version 2 - Activities

C.1. Experienced Student Activities

Activities

INSTRUCTIONS

The goal of this practical experimentation is to investigate text statistics and term weights from a pre-existing document collection and compare the results with the theory covered in lectures 2 and 5. For this exercise a tool is provided. In the tool select the collection “Cranfield 100” if you have an even student ID number and Cranfield “Collection 50” if you have an odd student ID number. Administer the following tasks and document your findings in a report.

Access to the tool:

Link: <https://infret.me>

Username: tugraz

Password: isr2017

TASKS

1. Text statistics

Administer the probability distribution on the level of single characters (Statistics → Letter distribution). Provide your findings in a table. Discuss the findings and find 2 application scenarios for information retrieval making use of the letter distribution.

2. Text statistics

Investigate the word distribution (Statistics → Word distribution) of the first 40 words (Rank, Word count, Percentage). Document the findings in a table. Show in a graph how “ $r * p_r$ ” varies (X axis should be the word and Y axis should be

the $r * p_r$).

Apply the following formula and find for the factor “alpha” the best value (least square sum error).

$$f_r = f(r) = f_1 r^{-\alpha}$$

Apply the following formula and discuss the results. Find the best value for “A” (least square sum error).

$$p_r = \frac{f_r}{f} = \frac{A}{r}$$

3. Distribution of words in documents

Investigate the distribution of words in documents (Weighting → Term Frequency weights → Raw frequency). Focus on the documents with IDs from 40001 to 40005 for Cranfield 50 and IDs from 10001 to 10005 for Cranfield 100. Compare the word frequency of the words “the”, “of”, “and”, “a”, “in” for the above mentioned documents. Provide your findings in a table.

Moodle Questions

- For Cranfield 50:
 1. How many times does the stop-word “in” occur in the first 5 documents? [20, 22, 25]
- For Cranfield 100:
 1. How many times does the stop-word “of” occur in the first 5 documents? [20, 22, 25]

Remove the stop words and compare the frequencies of the words “flow”, “boundary”, “layer”, “heat” and “theory” for Cranfield 50 and words “flow”, “boundary”, “layer”, “pressure” and “temperature” for Cranfield 100 for the above mentioned documents. Provide your findings in a table.

Moodle Questions

- For Cranfield 50:
 1. Which non stop word is the 4th most often occurring word in the document collection? [boundary, pressure, flow, heat]
 2. How many times does that word occur in the first 5 documents? [1, 6, 9]
- For Cranfield 100:

C.1. Experienced Student Activities

- 1. Which non stop word is the 4th most often occurring word in the document collection? [boundary, pressure, flow, heat]
- 2. How many times does that word occur in the first 5 documents? [1, 6, 9]

Compare the number of words left after removing stop words (every page except the last one has 50 words). What do you notice? Write your findings in the report.

Moodle Questions

- For Cranfield 50:
 1. How many stop words are there in the 50 documents? [95, 107, 116]
- For Cranfield 100:
 1. How many stop words are there in the 100 documents? [95, 107, 116]

[HINT 1: Use the basic selection data and the page numbers to see the number of words, each page except the last one has 50 words]

[HINT 2: All terms in the heat map are sorted in descending order by term frequency]

4. Inverse document frequency

Investigate the inverse document frequencies of words (Weighting → Inverse document frequency → Inverse frequency). Provide the IDF value, Frequency, Number of documents the word appears in for the words “flow”, “transition” and “flutter” in a table.

Moodle Questions

- For Cranfield 50:
 1. What is the IDF of flutter?
- For Cranfield 100:
 1. What is the IDF of flutter?

Compare the IDFs of words to their TF weights (weighting → Term frequency weights → Log normalization). Provide the Number of Term frequency weights larger than 3 (TF weight > 3) for words “flow”, “transition” and “flutter” in the same table. What connections do you notice between them and the IDF values? Write your answer in the report.

Moodle Questions

- For Cranfield 50:
 1. Which of the 3 words is relevant for the most documents based on the count of high TF values ($TF > 3$)? [flow, transition, flutter]
 2. Observe the term frequency and the TF weight values for the three words. What kind of correlation do they have? [positive, negative]
 3. Observe the term frequency and the IDF values for the three words. What kind of correlation do they have? [positive, negative]
- For Cranfield 100:
 1. Which of the 3 words is relevant for most documents based on the count of high TF weight values ($TF \text{ weight} > 3$)? [flow, transition, flutter]
 2. Observe the term frequency and the TF weight values for the three words. What kind of correlation do they have? [positive, negative]
 3. Observe the term frequency and the IDF values for the three words. What kind of correlation do they have? [positive, negative]

[HINT 1: Increase the min. Term frequency and remove stop words for easier navigation.]

[HINT 2: Use search inside the browser in order to find the relevant words faster.]

5. Term weighting

Investigate TF weight values (Weighting \rightarrow Term frequency weights \rightarrow Normalization max) for the words “the”, “a”, “wing”, “thermal”, “time” in documents with IDs from 40028 to 40032 for Cranfield 50 and from 10028 to 10032 for Cranfield 100. Provide the values in a table. How do the TF weights compare? Discuss the results in the report.

hl[HINT 1: All words except time can be found on the first page either when showing or hiding the stop words. Time can be found on the 4th page when hiding stop words for Cranfield 100 and 3rd page for Cranfield 50]

Moodle Questions

- For Cranfield 50:
 1. Select the two words that are more relevant for document 40030 according to TF. [the, a, wing, thermal, time]
- For Cranfield 100:

- 1. Select the two words that are more relevant for document 10030 according to TF. [the, a, wing, thermal, time]

Investigate TF-IDF weight values (Weighting \rightarrow TF-IDF \rightarrow TF-IDF weight 2) for the same words and the same documents. Provide the values in a table. How do the TF-IDF weights compare? Which weight do you find more insightful? Discuss the results in the report.

Moodle Questions

- For Cranfield 50:
 1. Select the two words that are the most relevant for document 40030 according to TF-IDF [the, a, wing, thermal, time]
- For Cranfield 100:
 1. Select the two words that are the most relevant for document 10029 according to TF-IDF [the, a, wing, thermal, time]

6. Reflection

Select the document collection "Fruit" and remove the stop words. Observe the weight values of the terms in the document collection when using the TF weight formula "Log normalization" (Weighting \rightarrow Term frequency weights \rightarrow Log normalization).

Moodle Questions

- Group 1
 1. What is the most relevant term for the document 60001 when using the TF weight formula "Log normalization"? [this, apple, shiny, computer, all terms are equally relevant]
- Group 2
 1. What is the most relevant term for the document 60003 when using the TF weight formula "Log normalization"? [this, apple, shiny, computer, none of the above]

Observe the weight values when using the TF-IDF weight formula "TF-IDF weight 2" (Weighting \rightarrow TF-IDF \rightarrow TF-IDF weight 2).

Moodle Questions

- Group 1
 1. What is the most relevant term for the document 60003 when using the TF-IDF weight formula “TF-IDF weight 2” [this, apple, shiny, computer, none of the listed]
- Group 2
 1. What is the most relevant term for the document 60001 when using the TF-IDF weight formula “TF-IDF weight 2”? [this, apple, shiny, computer, none of the listed]

Compare the TF weight values and the IDF weight values. Write your observations down in the report.

Moodle Questions

- Group 1
 1. Terms occurring rarely in a document: [will have a high TF weight value, will have a low TF weight value, will not have their TF weight values influenced by their occurrence]
 2. Terms occurring often in a document collection and rarely in a document: [will have a high TF-IDF weight value for that document, will have a low TF-IDF weight value for that document, will not have their TF-IDF weight values influenced by their occurrence]
- Group 2
 1. Terms occurring often in a document: [will have a high TF weight value, will have a low TF weight value, will not have their TF weight values influenced by their occurrence]
 2. Terms occurring rarely in a document collection and often in a document: [will have a high TF-IDF weight value for that document, will have a low TF-IDF weight value for that document, will not have their TF-IDF weight values influenced by their occurrence]

C.2. Novice Student Activities

Understanding Text Statistics and Term Weighting

INSTRUCTIONS

The goal of this practical experimentation is to investigate text statistics and term weights from a pre-existing document collection. For this exercise a tool is provided. In the tool select the collection “Cranfield 100”, administer the

following tasks and answer the questions.

Access to the tool:

Link: <https://infret.me>

Username: tugraz

Password: isr2017

[HINT] In order to compare the different values it might be easier to first copy the values in a table such as excel or google sheet.

TASKS

1. Text statistics

Administer the probability distribution on the level of single characters (Statistics → Letter distribution). Discuss the findings and find possible application scenarios for data and information retrieval making use of the letter distribution.

2. Text statistics

Investigate the number of distinct words (Statistics → Number of distinct words) for your text collection. Compare the number of distinct words in vocabulary with the number of words in the entire collection.

Questions

1. What do you notice?

3. Distribution of words in documents

Investigate the distribution of words in documents (Weighting → Raw frequency). Focus on the documents with IDs from 10001 to 10005. Compare the word frequency of the words “the”, “of”, “and”, “a”, “in” for the above mentioned documents.

Questions

1. How many times does the stop-word “of” occur in the first 5 documents?
[20, 22, 25]

Remove the stop words and compare the frequencies of the words “flow”, “boundary”, “layer”, “pressure” and “temperature” for the above mentioned documents.

Questions

1. Which non stop word is the 4th most often occurring word in the document collection? [boundary, pressure, flow, heat]
2. How many times does that word occur in the first 5 documents? [1, 6, 9]

Compare the number of words left after removing stop words (every page except the last one has 50 words). What do you notice?

Questions

1. How many stop words are there in the 100 documents? [95, 107, 116]

[HINT 1] Use the basic selection data and the page numbers to see the number of words, each page except the last one has 50 words.

[HINT 2] All terms in the heat map are sorted in descending order by term frequency.

4. Inverse document frequency

Investigate the inverse document frequencies (IDF) of words (Weighting → Inverse document frequency → Inverse frequency). Investigate the IDF value, Frequency, Number of documents the word appears in for the words “flow”, “transition” and “flutter”.

Questions

1. What is the IDF of flutter?

Compare the IDFs of words to their term frequency weights (weighting → Term frequency weights → Log normalization). Observe the Number of Term frequency weights larger than 3 (TF weight > 3) for words “flow”, “transition” and “flutter”.

Questions

1. What connections do you notice between the number of term frequency weights larger than 3 and the IDF values?
2. Which of the 3 words is relevant for most documents based on the count of high term frequency weight values (TF weight > 3)? [flow, transition, flutter]

3. Observe the term frequency and the TF weight values for the three words. What kind of correlation do they have? [positive, negative]
4. Observe the term frequency and the IDF values for the three words. What kind of correlation do they have? [positive, negative]

[HINT 1] Increase the min. Term frequency and remove stop words for easier navigation.

[HINT 2] Use search inside the browser in order to find the relevant words faster.

[HINT 3] The word “flutter” can be found on the fourth page when the “Remove Stop Word” is active.

5. Term weighting

Investigate term frequency (TF) weight values (Weighting → Term frequency weights → Normalization max) for the words “the”, “a”, “wing”, “thermal”, “time” in documents with IDs from 10028 to 10032. [HINT 1] All words except

time can be found on the first page either when showing or hiding the stop words. The word “time” can be found on the 4th page when hiding stop words.

Questions

1. How do the TF weights compare?
2. Select the two words that are more relevant for document 10030 according to TF. [the, a, wing, thermal, time]

Investigate term frequency - inverse document frequency (TF-IDF) weight values (Weighting → TF-IDF → TF-IDF weight 2) for the same words and the same documents.

Questions

1. How do the TF-IDF weights compare?
2. Which weight do you find more insightful?
3. Select the two words that are the most relevant for document 10029 according to TF-IDF [the, a, wing, thermal, time]

6. Reflection

Select the document collection “Fruit” and remove the stop words. Observe the weight values of the terms in the document collection when using the TF

Appendix C. Infret Prototype Version 2 - Activities

weight formula “Log normalization” (Weighting → Term frequency weights → Log normalization).

1. What is the most relevant term for the document 60003 when using the TF weight formula “Log normalization”? [this, apple, shiny, computer, none of the above]

Observe the weight values when using the TF-IDF weight formula “TF-IDF weight 2” (Weighting → TF-IDF → TF-IDF weight 2).

Questions

1. What is the most relevant term for the document 60001 when using the TF-IDF weight formula “TF-IDF weight 2”? [this, apple, shiny, computer, none of the above]

Compare the TF weight values and the IDF weight values.

Questions

1. Terms occurring often in a document: [will have a high TF weight value, will have a low TF weight value, will not have their TF weight values influenced by their occurrence]
2. Terms occurring often in a document: [will have a high TF-IDF weight value for that document, will have a low TF-IDF weight value for that document, will not have their TF-IDF weight values influenced by their occurrence]

Appendix D.

Infret Prototype Version 2 - Survey

Thank you for participating in this survey. This survey is being conducted to optimize the experimentation tool and the learning model for further courses. The survey results are anonymised and will be solely used for evaluation and research purposes. Your participation is completely voluntary. Completion of this survey implies your consent for the purposes stated above.

Please specify your gender

- Female
- Male
- Non-binary
- Other gender not listed, please specify
- Prefer not to say

Please specify your age

- 12-17 years
- 18-24 years
- 25-34 years
- 35-44 years
- 45-54 years
- 55-64 years
- 65-74 years
- 75 years or older

Please select your study program

Study programs listed either for TU Graz students or for RMIT students.

In case of other please write down the study program

Free form entry field.

Please select the number of semesters you have attended in your current program

Number entry field.

Add explanation if necessary

Free form entry field.

Are you a full time or a part-time student

- Full time
- Part time
- Other

In case of other please write down the answer

Free form entry field.

Prior knowledge

Which statement best describes your prior knowledge of text statistics?

- Can explain what they are and used them before
- Can explain what they are but never used them
- I have an idea of what they are but don't know where or when to use them
- Never heard of them or heard of them but don't know what they are

Which statement best describes your prior knowledge of term weighting approaches?

- Can explain what they are and used them before
- Can explain what they are but never used them
- I have an idea of what they are but don't know where or when to use them
- Never heard of them or heard of them but don't know what they are

Which statement best describes your prior knowledge of basic statistics concepts (probability, standard deviation, probability distributions, ...)?

- Can explain what they are and used them before
- Can explain what they are but never used them
- I have an idea of what they are but don't know where or when to use them

- Never heard of them or heard of them but don't know what they are

Which statement best describes your prior experience with basic data analysis approaches (observing data distribution, looking for correlation, ...)?

- Can explain what they are and used them before
- Can explain what they are but never used them
- I have an idea of what they are but don't know where or when to use them
- Never heard of them or heard of them but don't know what they are

E-learning tools help me learn course content.

Likert scale.

Which statement best describes your prior experience with using interactive simulations to learn concepts?

- Used them in multiple classes
- Used them in one or two classes
- Never used them
- Never used them and don't know what they are

Evaluation

General

It is easier to grasp the concepts by first exploring the data and not having complete knowledge of the theory from the start.

Likert scale.

It is easier to grasp the concepts by having complete knowledge of the theory and then exploring the data.

Likert scale.

Text statistics

This tool has helped me to better understand and reflect aspects on text statistics.

Likert scale.

The bar chart visualisation helped me understand text statistics.

Likert scale.

The line chart visualisation helped me understand text statistics.

Likert scale.

What parts of the tool helped you / didn't help you especially with the understanding of text statistics?

Free form entry field.

How did / didn't the bar chart visualisation help you understand text statistics?

Free form entry field.

How did / didn't the bar chart visualisation help you understand text statistics?

Free form entry field.

Term weighting

This tool has helped me to better understand and reflect aspects on TF weights.

Likert scale.

This tool has helped me to better understand and reflect aspects on IDF.

Likert scale.

This tool has helped me to better understand and reflect aspects on TF-IDF weights.

Likert scale.

The bar chart visualisation helped me understand idf.

Likert scale.

This tool has helped me to better understand and reflect aspects on TF-IDF weights.

Likert scale.

The bar chart visualisation helped me understand idf.

Likert scale.

The heat-map helped me understand term weighting.

Likert scale.

The formula explanations when clicking the help button helped me understand the theory better than just exploring the data.

Likert scale.

What parts of the tool helped you / didn't help you especially with the understanding of term frequency?

Free form entry field.

What parts of the tool helped you / didn't help you especially with the understanding of IDF?

Free form entry field.

What parts of the tool helped you / didn't help you especially with the understanding of TF-IDF?

Free form entry field.

How did / didn't the bar chart visualisation help you understand idf?

Free form entry field.

How did the heat-map help you understand term weighting?

Free form entry field.

How did the formula explanations help you understand the theory better than just exploring the data?

Free form entry field.

Usability questions (SUS)

All questions have Likert scale answer possibilities.

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

Computer Emotion Scale (CES)

All questions have the following answer possibilities:

- None of the time
 - Some of the time
 - Most of the time
 - All of the time
-
- Using this software, I feel Satisfied
 - Using this software, I feel Disheartened
 - Using this software, I feel Anxious
 - Using this software, I feel Irritable
 - Using this software, I feel Excited
 - Using this software, I feel Dispirited
 - Using this software, I feel Insecure
 - Using this software, I feel Frustrated
 - Using this software, I feel Curious
 - Using this software, I feel Helpless
 - Using this software, I feel Nervous
 - Using this software, I feel Angry

Questions to tie everything together

I find the tool useful.

Likert scale.

This tool makes learning the theory more interesting.

Likert scale.

After using this tool I feel curious about other information retrieval concepts that I haven't learned yet.

Likert scale.

After using this tool I would like to learn more about information retrieval.

Likert scale.

I like experiments and hands-on activities.

Likert scale.

Generally, experiments and hands-on activities help me to better understand theory and methods.

Likert scale.

Coding the IR concepts (term weighting, text statistics, ...) would have helped me to better understand the theory and methods.

Likert scale.

I would like to use such a tool for other subjects in ISR.

Likert scale.

For which other areas of ISR would you like to use this tool?

Free form entry field.

For which other subjects could such an approach be used?

Free form entry field.

Exploring the formulas using different visualisations helped me understand the theory more than just calculating them on paper.

Likert scale.

Appendix D. Infret Prototype Version 2 - Survey

What did you like the most?

Free form entry field.

What did you not like at all?

Free form entry field.

What could be improved?

Free form entry field.

Appendix E.

UI Guideline

Appendix E. UI Guideline

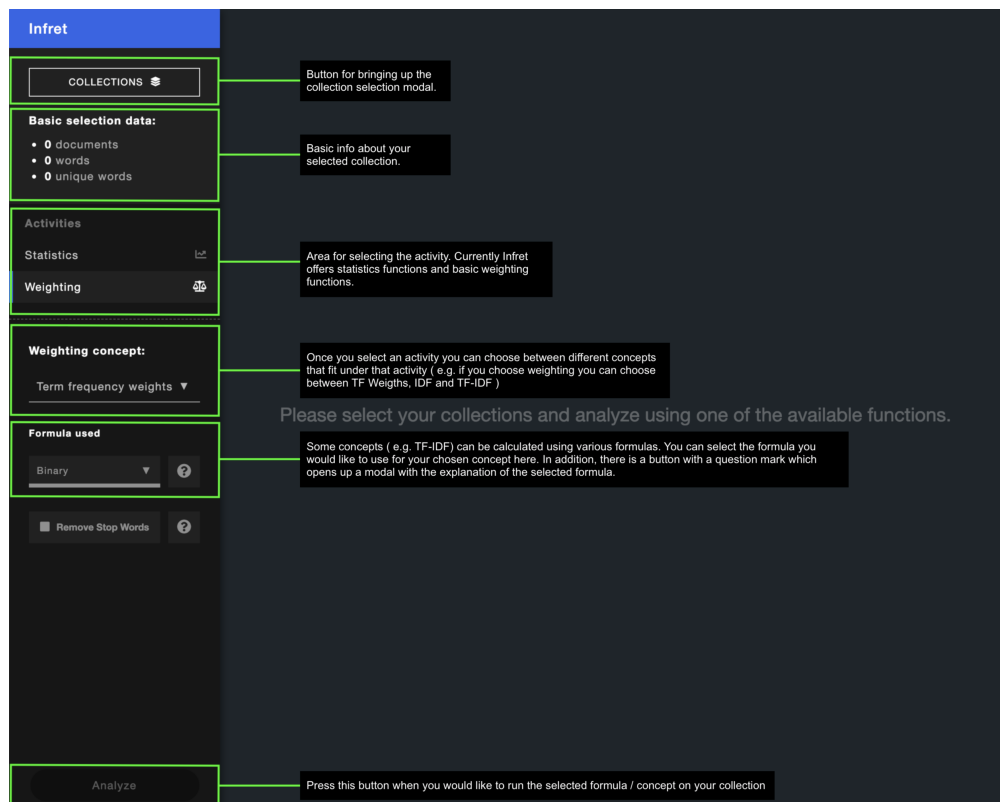


Figure E.1.: The first picture of the UI guideline handed out to experienced students.

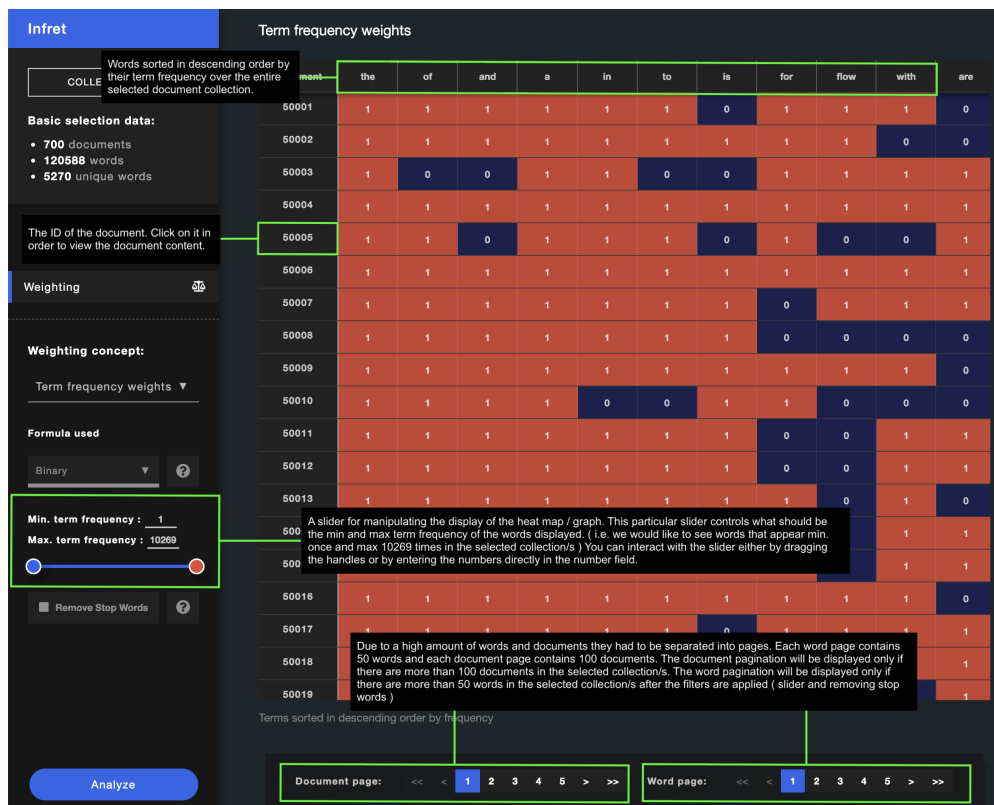


Figure E.2.: The second picture of the UI guideline handed out to experienced students.