HTW BERLIN - UNIVERSITY OF APPLIED SCIENCES

Bachelor's Thesis

Facilitating the use of Wikidata in Wikimedia projects with a user-centered design approach

by

Charlene Kritschmar

03. March 2016

Matriculation Number: 537961

Faculty IV: School of IT, Communication, and Business

Program of Study: International Media & Computing

First Examiner: Prof. Dr. Debora Weber-Wulff

Second Examiner: Dr. Michael Minge

Abstract

Facilitating the use of Wikidata in Wikimedia projects with a user-centered design approach

by Charlene Kritschmar

Wikidata is a Wikimedia project which stores semantical data in its knowledgebase, to be used by other Wikimedia projects. In its current form, the data is not used to its full potential for a multitude of reasons, as user acceptance is low and the process of data integration is unintuitive and complicated for users.

This thesis aims to develop a concept using user-centered design to facilitate the editing of Wikidata data from Wikipedia. With the involvement of the Wikimedia community, a system is designed which integrates with pre-existing work flows. The implementation makes use of various tools and design methodologies from the field of Usability Engineering, with which a solution is created that satisfies different types of users in regards to their past volunteering experience for Wikimedia projects.

The thesis concludes with a solution which may serve as a basis which may be implemented into various Wiki projects in the future and thus help make the editing of Wikidata from within another Wikimedia project more user-friendly and improve its acceptance in the community.

Acknowledgements

Contents

Appendix

A Persona: Miriam Weiland

B Persona: Sam Doughty

C Persona: Sven Kueppner

D Persona: Peter Wellington

E Persona: Tanja Brinks

F Persona: Jamie Morgenstern

G Persona: Max Senkmann

H Usability Test Evaluation

I Questionnaire

J Self-Assessment

Declaration of Authorship

Chapter 1

Introduction

1.1   Motivation

Wikidata, is a project by the Wikimedia Foundation [40]. It is a free linked knowledge-base. The Wikimedia Foundation was founded in 2003 as a non-profit organization whose goal it is to create and support open content to the public for free [19]. Launched in 2012, it aims to support its Wikimedia sister projects, for example Wikipedia and third-party projects, like Histropedia, a tool to visualize anything on a timeline, or Platypus, an open question answering tool.

A machine and human readable knowledge base means that much of the data used throughout the different Wikipedias could be stored in a central location and accessed from there, language-independent and free [40]. Especially the generation of infoboxes —structured metadata summarizing the respective article—would profit immensely due to the fact that they consist of pure data and no continuous text.

To illustrate with an example, one can look at the population of Berlin throughout various language Wikipedias. The German Wikipedia will be updating the number quite regularly, while other Wikipedias might use older numbers, creating inconsistent and outdated pieces of information. Even when comparing a selected *featured* articles in other languages —which are articles that have been promoted for their high quality by Wikipedia editors—on Berlin to the German article, it quickly becomes apparent that almost all values for the Berlin population differ from each other. This problem could easily be solved by importing such data from Wikidata, making the available information in infoboxes consistent throughout the different language sites. Populating the infobox on Wikipedia with data from Wikidata was a main goal of the creators[25, 27].

Unfortunately, at this time, the utilization of the existing data is very minimal. Roughly *three million* of the articles in the English Wikipedia and almost *half a million* in the German Wikipedia use data from Wikidata in one form or another (for example in the infobox or for geo-coordinates)[41]. These numbers are far below its potential when compared to the amount of articles the English and German Wikipedia have, that could make use of the data.

The lack of use has many reasons, such as being too complicated or not well enough documented, which will be discussed in the following section. This thesis aims to create, explain and implement a concept and design of infoboxes whose goal it is to increase the usage of the data in Wikidata. This is achieved with a user-centered design approach, since ultimately, the local Wikimedia communities will be using it.

1.2   Initial Problem Statement

Currently, the editing process involves being directed away from the *client* (unless otherwise specified *client* is referring to Wikipedia), Wikipedia to Wikidata, and, after submitting changes to a page, being redirected from Wikidata back to the client. This makes many users uncomfortable as Wikipedia editors are being forced to leave their local Wiki [26].

One also has to take into consideration that many Wikipedians do not know what Wikidata is, much less how to make use of it. This goes hand in hand with the ongoing issue of "not my Wiki!" [44]. It is often the case that Wikipedians as well as people from other Wikis prefer not to leave their local Wiki to avoid outsourcing information. Similar events have occurred when Wikimedia Commons was launched in 2004. It is an free-use repository for all types of media files, such as images for example, and it used used across all other Wikimedia projects [5].

Many Wikipedians initially refused to upload media files to Wikimedia Commons, due to a lack of understanding the advantages which a common media repository brings, amongst other concerns. However, eventually the popularity of the project increased to the extent that it is now widely accepted in the Wikipedia and other Wikimedia communities [10].

1.3 Objective targets

To improve the current editing situation, a concept is needed which will ease the editing process for the editors when working from a client such as Wikipedia. The field of user-centered design yields many methods and tools to approach this problem. In the scope of this thesis, some will be tested and used to narrow down the possibilities to find an optimal solution.

Furthermore, if the community were not to be involved in the concept development, it can be assumed that an implementation of any kind would not be successful. This can be expected because of the fact that the community does not only represent the main user group but also have a vital role in the development process of all Wikimedia projects, due to its Web 2.0 nature.

Chapter 2

Fundamentals

2.1 Wikipedia

Wikipedia is a Wikimedia Foundation project which was initially created in 2001 to serve as a free and multilingual online-encyclopedia. The community working and maintaining the project consists almost entirely of volunteers. Anyone can contribute to its content anonymously, or pseudonymously [15]. As of January 2016, across all language Wikipedias, there are more than 38 million articles maintained by about 75,000 active (more than 5 edits per month) contributors [43]. Writing and editing articles has to follow certain guidelines that the respective language community agrees upon. Changes are made after discussions are being held for a certain period of time, where anyone can voice their opinion. Since opinions vary, there are often large disputes about what belongs in an article, and what does not[9].

All the editing work is written in *wiki markup*, also known as *wikitext language* and *wikicode*. It is a markup language that is converted by Wiki software to HTML. Most language Wikipedias offer the use of a *WYSIWYG (What You See Is What You Get)* editor called *VisualEditor*. It lets users directly see the changes they made to an article and does not require any knowledge of wiki markup.

2.1.1 Infobox

Infoboxes are an optional source of information displayed next to the text of an article. They are exclusively located in the top right corner, for all left-to-right languages, and are commonly based on a template the community agreed upon beforehand. Infoboxes hold facts and statistics [8] about the article in a mostly machine-readable format. They function as fact-sheets to provide an overview about the subject of an article so the reader can quickly compare and grasp what the article is about. The values of the fields are written in wikitext, however, the Infobox itself is pre-built, thus it is not a requirement to know the necessary fields for the current article when creating an infobox, just the correct template suffices.

The exemplary article used throughout this thesis will be *Douglas Adams*. The infobox template used *Infobox:Writer*.
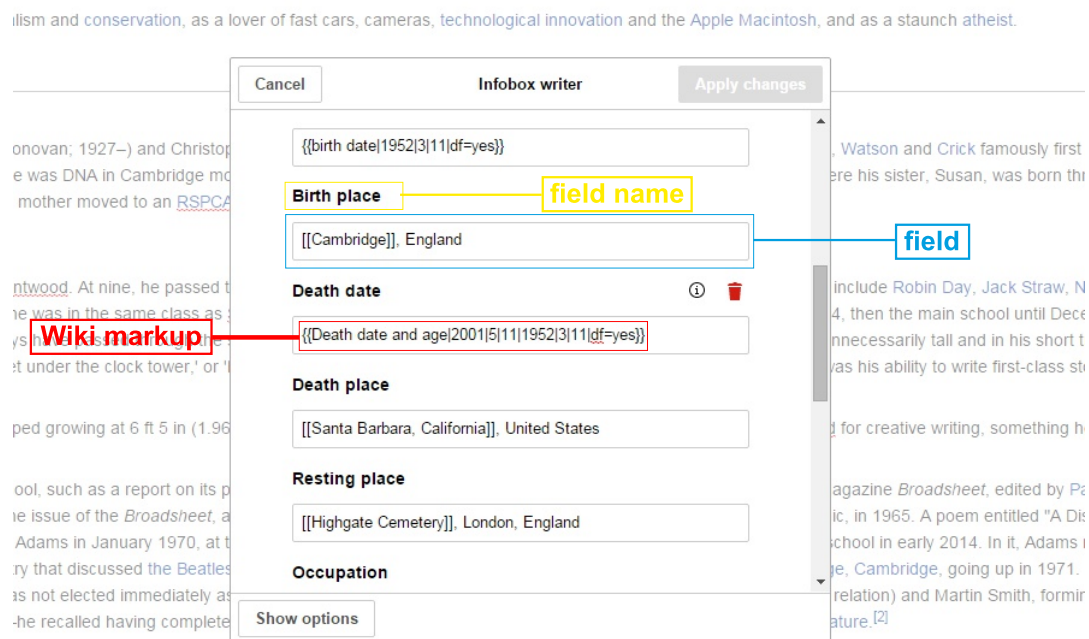
FIGURE 2.1: *Infobox*: Representation in VisualEditor on Wikipedia

2.1.2 VisualEditor

Unless otherwise specified, *editing* refers to editing with VisualEditor throughout this thesis. Firstly, it is important to mention that VisualEditor is not available by default on all Wikipedias. However, it is available as an experimental beta feature on all Wikipedias with the only exception being the English Wikipedia, where it is enabled by default [16].

The VisualEditor is an editing tool that lets users make edits through a graphical user interface (GUI), rather than constructing the complete article in wikitext. When editing an infobox, the expected fields are provided, thus an editor does not have to look them up and write them down manually. The fields are provided by the infobox templates which consist of information the community agreed upon as shown in this template documentation for a person [13, 16].

2.1.3 Community

In the environment surrounding the Wikimedia project, the word *Community* refers to the many people who volunteer their free time to maintain and create content in their local Wikis or contribute in another way, for example writing software [17]. As mentioned before, the community of the English Wikipedia consists of around 31 000 active contributors, the German Wikipedia has around 6000 active editors averaged over the year 2015 [42]. An editor counts as *active* when they contribute 5 or more times a month.

It is important to understand the structure of the community as this affects how much influence editors have over articles but also changes in the design. Within the communities there are different roles each member can have within a hierarchical structure. A new user without an account is called *unregistered user* and occupies the lowest position in the structure. Therefore, they have the least rights. The highest position a member can occupy is the one of a *Steward*. In between these two extremes are regular *registered users*, *administrators*, *bureaucrats* and many more. Each role grants the user special rights, but also brings certain responsibilities with it [14, 17].

2.2 Wikidata

2.2.1 About Wikidata

Launched in 2012, Wikidata has been providing a free linked database of structured data [40]. Structured data is machine readable and usually stored in a relational database [6]. All data is community-generated, maintained and published under a no-attribution CC-0 license, waiving all rights making their data public domain [1]. It is inherently multilingual and thus can be used across many platforms. Similar to Wikipedia, Wikidata can be edited anonymously or with a Wikimedia account [40].

As of February 2016, Wikidata has over 16 million entries[39]. The data is not as heavily utilized as hoped for by the organization and its community[41]. It has been hypothesized

that this is most likely caused by a multitude of reasons which possibly include the lack of usability of the data integration and editing functions and their visibility as well as a lack of trust in the quality of the data [21].

2.2.2 Structure

To understand how Wikidata works, it is helpful to take a look at the structure of an *item*. Items, in Wikidata, are representative of an entity which has properties. To give an example, an item could be anything from a specific person, to a concept or a place. Everything that has an entry in Wikipedia can also be an item in Wikidata. Every item on Wikidata has a human-readable name called *label* as well as a unique identifying number which follows a pattern of the letter Q followed by a numeric value (e.g. *Q12345*) [35]. Throughout this thesis the exemplary item of *Douglas Adams* (identifying number: *Q42*) is used.

Every item can also have one or more aliases, which represent additional names under which the item may also be known. In the case of *Q42* two aliases are set, one of which is *Douglas Noele Adams*, while the second alias *Douglas Noël Adams* is a similar variant, both including the middle name of the subject. A description can be added to every item, as in the case of *Adams*, it is *English writer and humorist* [35].
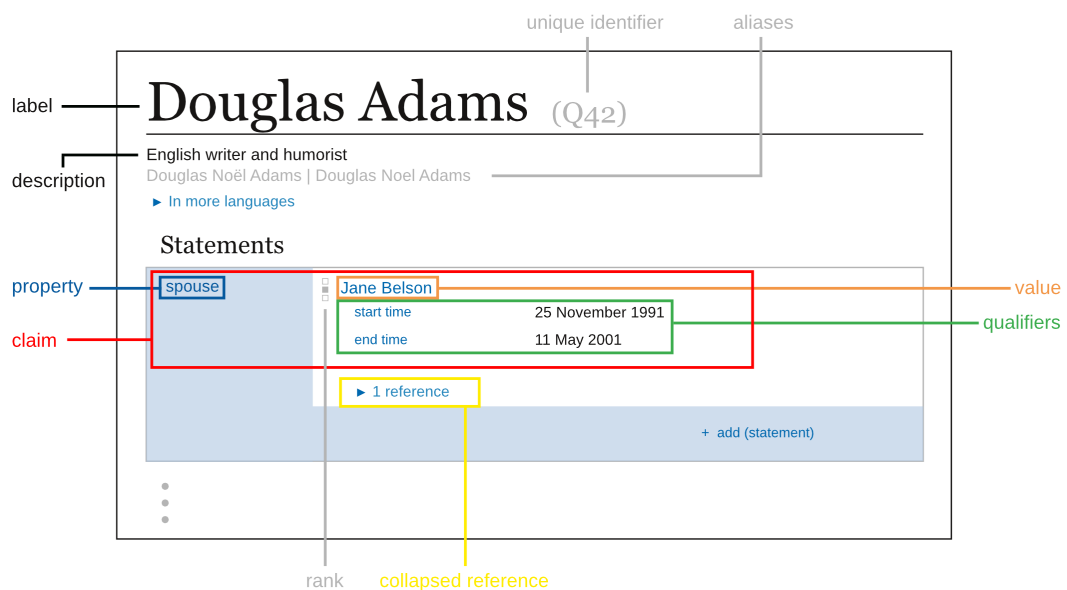

FIGURE 2.2: *Composition of an Item in Wikidata*: with missing sitelinks.

Every item can have a list of statements about that item. Statements encompass elements which describe details about the item. It does this with *properties* and *values* [38]. Properties can be seen as categories for the data values of statements [36]. For a person,

7

this could be, but is not limited to, gender, parents, first and last name, date of birth, sex, siblings, occupation and children. These properties are paired with a *value*. Values hold the exact data of the statements. To name an example, the statement with the property *last name* has *Adams* as a value. If we look at the property *spouse*, one can see that the value is *Jane Belson*. Values can have *qualifiers*, to expand or contextualize the statement, for example to specify when exactly the value was true [37]. To name an example, in the case of Adams' spouse it is a start time and an end time, since after Adams' death, Jane Belson was no longer his spouse.

Each statement can have multiple values. For example, one can have multiple spouses or multiple children, each with their own qualifiers and these do not have to be mutually exclusive, while each new property makes a new statement.

A property with a value and its qualifiers combined is referred to as a *claim*. Each of these claims should preferably have one or more references to back it up. These references are very important for Wikidata, since it is a platform for free knowledge specifically trying to avoid imposing *the one truth* on anyone. Unlike Wikipedia, where it must be agreed upon one *concensus* which will be published, in Wikidata multiple values for the same statement can exist side by side. An example would be the statement with the property *residence*. Adams has lived in many places throughout his life and having multiple values for residence is possible without being mutually exclusive.

When a statement has multiple values it is sensible to rank them, in order to show which one is the currently most agreed upon, for example. In the case of Adams the current residency would hold the primary rank, and all other residencies would hold the normal rank. In the case of multiple children, for example, all of the statements would be ranked 'normal' because one is not more correct than the other. The third available rank is 'deprecated' which is used for known wrong values. At this point, one can see how qualifiers and ranks play an important role in Wikidata. Qualifiers help to clarify when, where or under which circumstances a value of a statement is accurate and ranks help make sense of them in relation to each other.

An item additionally functions as a link between all Wikipedia articles on the same topic throughout the different languages. This means that an item is language-independent to a certain extent. All referenced pictures and all other references to one of the Wikimedia projects are also interlinked at Wikidata. This makes it a secondary knowledge base because it allows to store all data with a reference to an original source.

2.3 Previous work on editing Wikidata from a client

The Russian Wikipedia developed a so-called *gadget* to edit Wikidata from Wikipedia. It is not possible to import this data to use in the article but solely to edit the respective item of the article. Since the infoboxes on the Russian Wikipedia have been partially modified to use Wikidata in some places, the edits potentially have an impact on the Wikipedia article, but indirectly.

Analyzing the design, it becomes clear that the gadget is relatively user-unfriendly to users new to Wikidata. This is because there are many buttons, which all look very similar to each other. This makes none of them stand out which results in a lack of clues for the user, where to start or what to do at all. When looking at the dialog principles defined by the ISO 9421-110 the *self-descriptiveness* is being entirely overlooked. Self-descriptiveness is the ability of any element or an entity as a whole to describe itself to the user and giving them a sense of locality and possible functionalities [29]. When hovering over the button, one gets a short explanation of the functionality of the button. Unfortunately, this is not consistent throughout all buttons since some lack any kind of explanation. The same disregard of the dialog principle can be applied here.
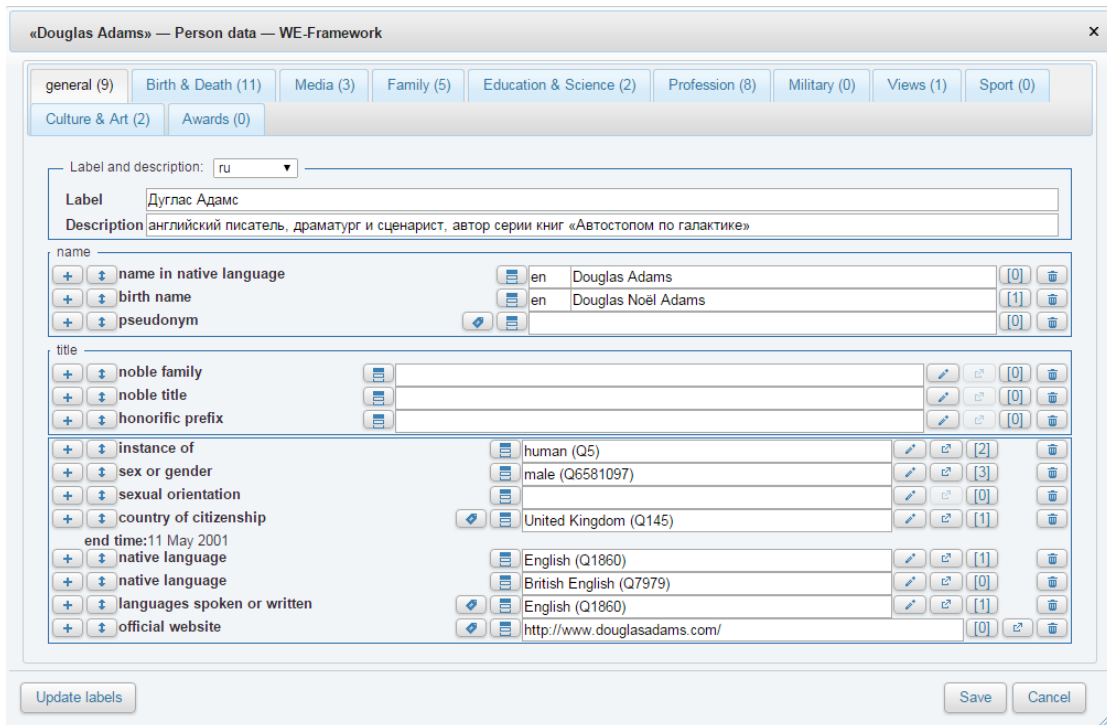

FIGURE 2.3: *Screenshot*: Wikidata editing gadget on the Russian Wikipedia

When hovering on the button to the immediate left of the text fields, the text says: *select snak type*. This could be a problem for inexperienced Wikidata users since they do not know what a *snak type* is and how to make use of it properly. Again, the self-descriptiveness is not sufficient.

Another example of bad design is the information the user gets when hovering over the double sided arrow button. This button has the functionality of changing the rank of the statement. This is not the symbol currently used on Wikidata for ranks, thus it would arguably even be confusing for experienced Wikidata editors. The dialog principle of *Conformity with user expectations* is not met here. This principle describes how a desired behavior should conform to agreed upon conventions [29]. However, the information it gives when being hovered over is only the current rank, which in most cases is *normal*. The user is not given any information as to what this is, and how to use it. Furthermore, it does not offer the possibility of getting more information on that functionality since the user is not given a name for this function.

Generally, one can say that the design elements are not being used efficiently to communicate helpful information to the user. It is important to mention that this gadget yields much editing power, meaning that all editing functionalities known from Wikidata directly, can be found here too. This explains—but does not justify—the design choices that were made. Since the goals of the tool developed by this thesis vary from the goals of the gadget used in the Russian Wikipedia, they can not be easily compared. The goals will be elaborated on in the following chapter.

The gadget is an interesting experiment and the first of its kind. There is more work to be done, in order to improve it, nevertheless it is a great first step to Wikidata more accessible and hopefully an inspiration to other language Wikis.

2.4 User-centered Design

User-centered design (UCD), also known as *human-centered design (HCD)*, is an approach to interactive systems development which aims to make systems usable and useful by focusing on the users, their needs and requirements, and by applying human factors/ergonomics, and usability knowledge and techniques." [30] Usability is defined by the *International Organization for Standardization* as the "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use." [28] *Effectiveness* is the "accuracy and completeness with which users achieve specified goals", *efficiency* is the "resources expended

in relation to the accuracy and completeness with which users achieve goals", and *satisfaction* is defined as the "freedom from discomfort, and positive attitudes towards the use of the product." [28]

Furthermore, UCD is founded on principles, which need to be followed in order to successfully implement the UCD-process. Thereby it is crucial for the designers to understand the users, their tasks and working environment. The users should be involved in the developmental process and product should be continuously adapted and improved on that basis. The process should be iterative and continuously take into account the whole user experience [18, p.109-110]. Usability should make a product easier to use while increasing productivity and decreasing training, cost and errors [23, p. 1-2].

The field encompassing usability, is called usability engineering. It is described as the framework to provide structured methods for achieving usability [23, p.2], rather than improving the specifics of design. The usability engineering lifecycle (UEL), as suggested by Mayhew in 1992 can be seen in her book *The Usability Engineering Lifecycle, 1999*. The UEL is made up of 3 main steps: Requirements analysis, design/testing/development, and installation. For this thesis, the requirements analysis and parts of the design/testing/development will be important. It is not always necessary to follow the exact steps described, but it is rather suggested to individually adapt the process to the project requirements.

The requirements analysis is accomplished by working through the following steps. Firstly, creating *user profiles*. This is possibly the most important step in user-centered design, since they will drive further decisions on design and functionality throughout the process. Furthermore user profiles are crucial when devising the next step, the *contextual task analysis*. In this step, the goals and needs of the users are acquired by observing or questioning users on their work-flow and problems. Next, the *platform capabilities and constraints* are determined. These will function as restrictions for feature implementations. Lastly, *general design principles* play an important role. They function similar to the platform capabilities in constraints, by setting restrictions on the design possibilities.

These four steps, user profiles, contextual task analysis, platform capabilities and general design principles, are all combined into forming the *usability goals*. These are qualitative and quantitative goals which are extracted from the four previous steps. These goals help focus the process in a desired direction, pushing it towards a more optimal outcome. The requirements analysis is the base for the whole design process. Everything that is developed later, especially in the first iteration which mainly encompasses this thesis, builds on the establishment of the requirements and its analysis.

As mentioned before, the second stage of the cycle, the design/testing/development stage, is the most flexible and should be molded according to the needs and the scope of the project. Generally, this part consists of three levels which are small iterations in themselves. Each level consists of several tasks of which the successful completion should be checked when the end of the level is reached. Only after the decision point at the end of a level is answered positively, one should proceed to the next level, otherwise one should re-iterate.

In the first level a conceptual model of the system is created, implemented to a prototype and evaluated. Level two deals with setting standards, and level three is based on the results of the first two levels [23, p.5-9]. For this project tasks from several levels were selected for the first iterations, since the scope of the thesis would only allow for one level. It should also be made clear, that this thesis in no way suggests, that one iterations is sufficient for a successful development of this product, but rather emphasizes that several more iterations are necessary to produce a successful result.

Following this process will help to create a product, which will put the needs and wishes of the users into focus, thus leading to higher efficiency and satisfaction for the users while decreasing frustration and errors.

Chapter 3

Requirements Analysis

3.1 Status Quo Analysis

Creating a concept for an editing tool requires an understanding of the editing process and the capabilities of the system it will be implemented on. This is necessary to develop a feasible and functional tool for most efficient usage.

All modifications to the data stored at Wikidata are done on Wikidata directly, by editors who are familiar with it. Edits are also made by bots, which are computer programs that fulfill certain tasks when executed, which are usually repetitive and too large in volume for humans to handle in a time-efficient manner [34]. Most editors and authors on Wikipedia are uncertain of the data quality on Wikidata and therefore have trouble trusting it [21]. The main arguments named in this ongoing discussion against Wikidata are the lack of referenced statements, and the low quality of the references for those which are referenced [21].

These arguments are not wrong, as a matter of fact, around half the statements on Wikidata are unreferenced, but this is a one-sided way of analyzing this issue. It should be taken in into account that Wikidata has, over the course of 2015, raised the amount of referenced statements from 12.7% to 20.9%. The previously made argument also does not take into account that many statements on Wikidata are self-evident and need no referencing. One such example are the links to images on Wikimedia Commons or the more than *three million* items with the statement *instance of: human* [26].

Nonetheless, these and other misconceptions exist in the minds of the Wikipedia editors and keep them from contributing to the project, and yet an increase in contribution is what Wikidata needs to be able to grow to its full potential according to the project manager of Wikidata, Lydia Pintscher [26].

For the editors that have no or almost no knowledge of Wikidata, it is not necessarily a project or concept they are willing to grasp. It will not be the task of the thesis to change their attitude towards Wikidata, but rather to implement the tool seamlessly with the already existing work-flow. The goal is for the aforementioned editors to not have to fully understand Wikidata, while nevertheless being able to contribute to it.

Presently, the extraction of information from an item to use in a Wikipedia article is only possible when knowing the exact Wiki text and syntax for implementing information from Wikidata. This leads to a less than average amount of usage of the items throughout the Wikimedia projects, since most editors do not know the specific rules of implementation, in addition to it being poorly documented, if available at all. Therefore, the items are rarely, if at all, used to populate the infobox or otherwise used in Wikipedia or other Wikimedia projects, which results in editors adding information manually.

Wikidata and Wikipedia are connected through so called *sitelinks*, that link an item on Wikidata to an article on Wikipedia throughout the different languages. When connected, one can access the Wikidata item directly through the respective article and vice versa. As a side effect, this sometimes leads to the discovery of duplicated articles when they are linked to the Wikidata item [12]. Despite this connection, the structural concepts of Wikidata and Wikipedia differ in many ways from each other, which will make a unification rather challenging, as well as making sure the editors have a correct understanding as well.

3.2 User Profiles

As mentioned in section 2.4, when designing a product or a tool in a user-centered design manner, the user is the focus point during the complete design process. Thus the first step in the design process is to determine the different users. This is usually done with a questionnaire, the consultation of an expert or a combination of both. With the pieces of information gathered from this approach, personas can be created and first requirements can be derived.

Personas are defined by Alan Cooper (1995) [2, p. 55] as *"not real people that are based on the behaviors and motivations of real people and represent them throughout the design process."* Personas are based on data, which is gathered through interviews and observation of actual users. When designing for a user group, personas serve as specific users that can be referred to, when thinking of their wishes and needs regarding the product in question.

3.2.1 Who are the users?

In order to get a valid representation of the user group, a question page was set up on the Wikidata project pages [7]. Project pages on Wikis function as a discussion page where anyone can discuss current projects and other specifics of the respective Wiki with other users. This means that anyone who is interested in the topic of this thesis can contribute their input. Naturally, this would not bring forth every type of user, but yield a good overview of users that the tool is targeting. The questions are also a meant to get an idea of the wishes of the users and their expectations. The categories for user input were:

- Current state and existing solutions

- What are the major problems you see right now with editing on a client?

- Do you have any general or specific ideas on how editing should work?

- Other input

After the helpful answers are filtered out, organized and broken down to their core message, following pieces of information could be learned from them:

- There is a similar implementation in *ruwiki*.

- A strong wish for it to be integrated with VisualEditor.

- There should be a focus on Wikipedia.

- Must not have a high learning curve like Wiki markup does.

- Must be careful not to facilitate vandalism.

- Restrict the editing power to a useful minimum.

- Include descriptions and helpful information on items, properties and their use.

- Automatic extraction from Wikidata to infobox.

- Make sure statements do not get overwritten but added.

- The templates of the infoboxes in Wikipedia will need to be edited.

The information gathered from the initial input will be crucial when creating the requirements for the tool, since they were extracted directly from potential user who have dealt with this subject to some degree.

As for narrowing down the user group, three types became apparent through the project page:

Firstly, users that would appreciate a tool to integrate data from Wikidata into their articles and are interested in its development.

Secondly, users that like Wikidata but do not want to invest time into learning something new and want the import of Wikidata's data in the infobox to be automated.

Lastly, very skeptical users that generally support Wikidata but are very suspicious of all changes and only look at the negative aspects.

It was fairly clear that the whole user spectrum is not included in these three types, so a common practice is to get expertise from someone who knows the community and its user very well [23, p.36]. In this case, a very good source of information is Lydia Pintscher, the Product and Community Manager for Wikidata at Wikimedia Germany e.V. After questioning her, it became apparent that three important users are missing. A Wikidata editor as well as a non-user trying to abuse the system.

3.2.2   Personas

Personas as a design tool, were introduced in 1995 by Alan Cooper in his Book *About Face: The essentials of interaction design.* There he describes personas acting as stand-ins for real users [2, p.59]. They are supposed to give the developers a clear picture on the needs and expectations of the user group, rather than representing specific users as well as help uncover the requirements. The benefit of using personas is that they give the user group a face and a name, they become relatable. They put the needs of the users into context by giving them a motivation and a background, and thus a driving behavior [2, 4, 31, p.59].

Personas can generally be separated into six different types of which five will be relevant to this thesis [2, p. 71-73]. *Primary* personas receive the most attention of the target group and cannot be satisfied by a design aimed at any other type of persona of the target group. Typically, if the primary personas needs are met, the other personas are satisfied to some extent or even fully. There can be multiple primary personas, whose needs all have to be met.

Secondary personas differ only marginally from the primary personas, by at the maximum a couple of diverging needs of the primary personas. Nevertheless they are too important to neglect.

Supplemental personas are the users that do not fit the primary or secondary personas, but would be satisfied by either the interface of the primary or secondary persona.

Served personas are not direct users, but profit from a good user interface. In this case this would be the Wikidata editors who would have to deal with an influx in vandalism if developed poorly.

The final type of personas, relevant for this thesis, are the *negative* personas. Negative Personas are not users of the product either, but rather serve the purpose for the developers to know who to not make the product for. In this case this would be users who aim to vandalize content on Wikidata through the tool.

An example of a primary persona which was created for this thesis:

Jeff Cohen

Demography:
- male
- 54 years old
- lives in seattle, washington
- married to Tom Cohen
- two daughters: Franscene and Rebecca
- science teacher at local highschool
- speaks english and dutch
- likes hiking and birdwatching
- organzies editing sessions in his community
- knows Wikidata but has a hard time using it

Behaviors:
- involved in multiple wiki projects
- admin/editor on a larger scale
- very community oriented
- likes to help people
- likes to organize things
- likes time to himself
- fatherly figure

Needs and goals:
- want to see the projects grow and improve
- expanding free knowledge
- help people have access to free knowledge
- concerned whether he's making a difference
- frustrated about the senseless vandalism
- fears that Wikipedia will stagnate at some point

"In en-wiki, there is no way any ordinary editor would have the slightest clue when creating an article that wikidata even exists and even less than the slightest clue how to create it. I've been editing nine years, nothing to help me other than talk page chitchat and completely incomprehensible instructions. Wiki markup syntax is complex enough for non-programmers, to go to a different web site and input something is too steep a learning curve for me..."

FIGURE 3.1: *Persona*: Makeup of a persona on the example of Jeff Cohen [33]

Jeff is 44 years old and works as a science teacher at a high school mainly teaching biology and chemistry. He has been teaching for almost 20 years and he still enjoys his job. He sometimes worries that the kids nowadays take the access to free knowledge for granted and thus wants to make sure that there will be enough future contributors out there. He runs a monthly open editing session to which he encourages his students to come as well. He has been involved in community work for a long time and is a respected member. He is not the most tech-savvy person but that has not stopped him from learning the Wiki markup language.

Fortunately, since December 2012 he can use VisualEditor which makes it much easier for him to maintain his pages. He loves the idea of Wikidata and would love to see the community use it more actively but he struggles with the usage and often can not be bothered to find out how to use it with the little instructions he can gather scattered around different pages. This results in him referring to the old ways. He would love a tool that is as easy to use as VisualEditor to integrate Wikidata into his infoboxes.

Altogether eight personas are used for the project of the thesis, of which 3 are primary personas. Their detailed information can be found in Appendices A - G. Following is a short overview of them:

The second primary persona which was developed for this thesis is *Miriam Weiland.* She is a highly specified editor though not very involved in the community and just wants to write her articles as efficiently and quickly as possible. Miriam would not be satisfied by the needs of Jeff, since she desires more automation than Jeff, and values efficiency over simplicity.

The last primary persona relevant to this thesis is *Sven Küppner.* He is very skeptical of outsourcing data to a different Wiki project and generally prefers to be in control of all edits made, much unlike Miriam.

Two secondary personas emerged in this thesis. The first one is *Sam Doughty.* He is a sporadic editor who is is almost entirely satisfied by the needs of the primary personas. Since he goes long periods without editing, he would need extra reminders or introductions to the tool unlike the frequent editors before him.

The second one is *Jamie Morgenstern.* She just started editing in general and is not very tech-savy as it is. She will need all the extra help to understand the tool and be able to use it appropriately.

Served personas include *Peter Wellington* and *Tanja Brinks.* Peter is a template creator on Wikipedia and would need to make adjustments to the templates, so the new functionality could be integrated into the infoboxes. Tanja is a Wikidata editor, whose main priority is to keep the Wikidata vandalism-free and raising the quality of the data.

The negative persona included in this thesis is *Max Senkmann.*

Further on, all the above mentioned personas are mapped on a spectrum in different categories, such as their experience in editing Wikipedia or their level of community engagement. This is done to make sure there are no major areas in the user group which were not covered as identifying potential duplicate personas [20].

3.2.3   User Scenarios

After creating every persona, user scenarios for each of them are developed based on the requirements extrapolated from each user. User scenarios are written under the assumption that the product already has all necessary functionalities and does what is expected of it. It includes details of the system and how the users might experience it. By engaging the user with the product, it helps to put them into a context with the product. [2, 3, 32, P. 80].

- Right after having dinner, Jeff decides to do some editing on Wikipedia. He boots up his computer and logs into Wikipedia with his account. There were some new discoveries in Ornithology and he would like to adjust the information in the respective articles. A new common ancestor was found and it turns out that there is a whole other new genus which requires an article of its own.

- Jeff prefers VisualEditor for editing. He long switched to using Wikidata for his articles because it saves him time and trouble. He calls up the article of the bird he wants to edit and clicks on VisualEditor. It boots up and he clicks on the infobox which then opens the editor. He scrolls down to the part where the genus can be edited and sees the little symbol signaling him that this information is coming form the respective Wikidata item. He can even hover over it to have this information displayed to him.

  He then is faced with two boxes asking him whether the information displayed has been correct at any point in time or whether it is just entirely wrong. He choses the first one because until the discovery of the new shared ancestor it did belong to the former genus. He then is asked to type in the name of the new genus and, if known, a date when it became the new accurate term. He is also asked to submit a reference.

  In this field he also has the option of looking at the entire list of values for that specific property. After confirming, he is informed that he added a value to the respective property. When he looks at the infobox the correct genus is being displayed.

- He checks if there is already an article for the new found common ancestor and finds out that there is not. He decides to start writing it, first writing all the text for the article, and then proceeding with the infobox. When he types in the name of the article it automatically suggests to use the corresponding Wikidata item instead.

  In this case, there is no item yet so it suggests to create one. He chooses to do so and a separate box opens up for him to add properties to the newly created

item. For now these are restricted to the infobox templates fields. All he needs to do, is fill out the fields as he would in VisualEditor and press ok. They are then automatically transferred and at the same time the item gets created, so the article can refer to the label's unique identifier.

- Jeff decides that he wants to add another property which is not part of the template. He clicks on the usual button at the bottom of VisualEditor to add another one. He types in the property's name and is automatically asked if he would like to add it as a value for the Wikidata item. He chooses yes and fills in the value for it.

- Because he used the Wikidata item in his infobox and text the item and the article are now automatically linked to each other. If anyone wants to create an article about this new genus in another language they already have all the information available.

Based on what was learned about Jeff, the following expectations can be extrapolated for the project:

- keep vandalism low

- have a low entry barrier

- should be usable with the knowledge he already has

- does not want too much to change since it took him so long to learn everything

- should be instantly doable

- should be the same language

- showcase the connection with Wikidata

- explain what Wikidata is briefly

- easily understandable tutorials and explanations should be offered

- should clearly differentiate that this is Wikidata data

3.3   Requirements Analysis

3.3.1   Functional Requirements

A feature that is important for all personas are the tool tips, but especially for Jeff, Sam and Jamie since they either have trouble picking up on new functionalities, edit very

rarely, or are unexperienced. It is important to design the product in order for it to blend in seamlessly with VisualEditor and can be used intuitively, but nevertheless, all users would appreciate tips when first trying to integrate Wikidata in their infoboxes.

Another necessity is to be able to switch all fields to import data from Wikidata with one click, whereas editing the actual data should require extra steps in order to prevent vandalism. Statements that get edited by default should get added instead of being overwritten. This requirement is derived from Miriam and Sam since they most appreciate efficiency when editing, and do not require control as the other personas.

A requirement concerning all personas is the ability to only use Wikidata in selected fields. This is especially important for Sven, since it grants a higher level of control, unlike switching all fields at once.

Furthermore, certain functions should not be allowed like deprecating statements and deleting whole statements. These features should still only be possible on Wikidata directly in order to prevent harsh vandalism, which would result in a degradation of the data quality. This requirement is relevant to all personas, however this is especially important for Max and Tanja. It functions as a barrier to prevent the most severe vandalisms.

The adding and editing of references, qualifiers and ranks should be possible. This is a general requirement, needed for the functionality of the tool that is relevant to all personas.

3.3.2   Non-Functional Requirements

There are many non-functional requirements which need to be met in order to ensure a positive user-experience. The learning curve should be kept as low as possible, in order to animate users to continually utilize the new feature rather than avoiding it. The users already completed a steep learning curve learning mark up and may not be willing to invest time to familiarize themselves with a new work flow, meaning the entry barrier should be low. Although all personas would have a great advantage from this requirement being met, Jeff, Jamie and Sam would profit the most, since they would most likely be deterred from using it if the requirement is not met.

The previous requirement should be realized with the following requirement in mind. It is imperative to not facilitate intended or unintended vandalism. This is not only important for Tanja and Max, since the quality of the data depends on it. Differentiating between vandalism makes sense, since it can be committed purposefully and happen accidentally.

In the latter case the user may not be fully aware of the implications of their edits and thus cause destruction of data.

Hence, the consequences an edit on Wikidata will have must be made clear to the user, since an edit would cause an alteration in every place the same item is being used, unlike edits made locally in an article.

Another important issue is locality. Edits, that are made locally must also convey that, as well as edits that are being made remotely. Otherwise, the user can get confused about where the edits will occur. This is, as most other requirements, relevant for all personas.

It will also be important to convey the two separate features the extension brings. Firstly, it will be possible to import the data from Wikidata into the fields and then it can be edited or secondly, it can also be edited, without being used in the fields. These two features can be entirely separate from each other.

The usability of VisualEditor should not be impaired for users who wish to not utilize the new set of features, as existing editing processes must not be disrupted or interfered with. This must be especially ensured for skeptical users, such as Sven.

3.3.3   Prioritization of Requirements

To more effectively create a product in a strict time frame as it is the case for this thesis, it is helpful to prioritize the requirements, in order to know where to put the focus and make the most valuable progress. A common method in software development is the *MoSCoW Prioritisation*. MoSCoW stands for **M**ust have, **S**hould have, **C**ould have and **W**on't have [11]. It is important to keep in mind that this method was applied to the scope of the thesis, and not the actual first time deployment of the product. This means that the requirements were sorted according to what was possible to achieve in one iteration and not the whole design process.

MUST Must itself stands for Minimum Usable Subset. These are the requirements which make the product a viable solution in the first place. Without them it would be illegal, and/or unsafe to use. If deploying the product without having fulfilled the requirements in question would be pointless they must be classified as MUST requirements. All other requirements, even if they seem crucial on the surface, should be placed in one of the following categories.

SHOULD Unlike MUST requirements, SHOULD requirements are not vital to the viability of the product. Nonetheless SHOULD requirements would cause a significant

drop in user satisfaction and product effectiveness when not implemented. Therefore it should be avoided to neglect them, potentially forcing the users to do a workaround and thus being more error prone.

COULD The main difference between COULD and SHOULD requirements is that if left out, COULD requirements would have less of an impact on the overall satisfaction of the user and the product's effectiveness. They are also often being referred to as 'nice to have' requirements. If they do not get implemented then the overall success of the product would not be much affected. When the project runs into trouble and an unforeseen interruption occurs, COULD requirements are the first to be eliminated.

WON'T The requirements labeled WON'T are agreed upon to not be delivered. They are recorded for future deployments, and act as a means to keep the focus on the three requirements categories above.

Functional requirements:

> **MUST:** integration into VisualEditor, two distinct input methods, toggling each field separately
>
> **SHOULD:** only show statements that are in the template, no statement deletion, no statement deprecation, check if corresponding item is blocked for editing
>
> **COULD:** tooltips, auto completion, blinking dot in tool bar to draw attention to the new feature, editing should be deactivatable by admin if controversial topic
>
> **WON'T:** preview of changes

Non-Functional requirements:

- suppression of vandalism

- low entry barrier / shallow learning curve

- sense of locality

- sense of consequences

- sense of ownership

Chapter 4

Design and Prototyping

4.1 Design principles

After creating the Personas and evaluating their needs and basing the functional and non-functional requirements on this evaluative process, initial design concepts were sketched out. Since the mock-up has to seamlessly integrate into VisualEditor. Moreover, it was important to not overload the current design, in order for the new appearance to not be overwhelming for the users.

To start a design process, one needs design guidelines as mentioned by Mayhew in [23, p.7]. For this thesis the dialog principles laid out by *ISO 9421:110* [29] and *Nielsen's ten heuristics from 1995* [24] were used as a guideline.

As the name already suggests, the ten heuristics by Nielsen consist of ten separate principles. The dialog principles consist of seven dialogs. For the purpose of the thesis, both guidelines will be merged since their core messages overlap in most areas. The dialog principles have been added in brackets, to ensure separate referencing if necessary.

Visibility: (Self-descriptiveness)
> The system should always inform the user of the current occurrences though timely feedback for example.

Match between systems and the real world:
> The system should speak the language of the user and it should be using phrases and words the user would understand instead of technical terms.

User control freedom: (Controllability)
> Users should always be able to navigate the interface freely and be in control of

movement and speed. An important example would be to always include an *undo* button in case the user accidentally navigated to an unwanted place.

Consistency and standards: (Conformity with user expectations)
The way elements are presented, positioned and react should be done in a consistent manner. If they get introduced in one way, they should not change throughout the product.

Error prevention: (Error tolerance)
Instead of having good error messages, it is even better to not let the mistakes happen in the first place. This can be achieved by giving very clear instructions or restricting a moving further if the correct input was not given.

Recognition rather than recall:
The user should not be forced to remember details throughout multiple steps. Navigational tools should always be clearly visible.

Flexibility and efficiency of use: (Suitability for individualization)
Features for use optimization and maximizing efficiency should be available for advanced users. Though these should not be in plain sight, as to not overwhelm new users.

Aesthetic and minimalist design: (Suitability for the task)
Interfaces should not contain superfluous information, since this would distract from the relevant information and unnecessarily confuse the user.

Help users recognize, diagnose and recover from errors: (Encouragement of learning)

Error messages should never include codes and be written in a human readable format, clearly stating the nature of the error.

Help and documentation:
Help should always be searchable, easy to find and should be focused on potential user tasks and provide concrete steps to solving the issue.

4.2   Mock-Up Development

Initial quick sketches are made on paper, and then sketched out with the software *Pencil*, also known as *Pencil Prototype*, developed by *Evolus*. It is a free GUI-prototyping software usable across all platforms.

The nature of the problem asks for a solution with two visually distinct modes. Throughout this thesis, the two modes will be referred to as as Wikidata mode and Wikipedia mode, where the latter is the already known state which the editors are used to, where the input is all manual using Wiki text. Wikidata mode is the new additional mode, where the values comes directly from Wikidata. The actual editing of Wikidata is restricted to this mode only.

The locations which needs to represent mode most clearly are the *fields* of the infobox. Based on this insight, the first design idea makes switching between modes possible at the fields.
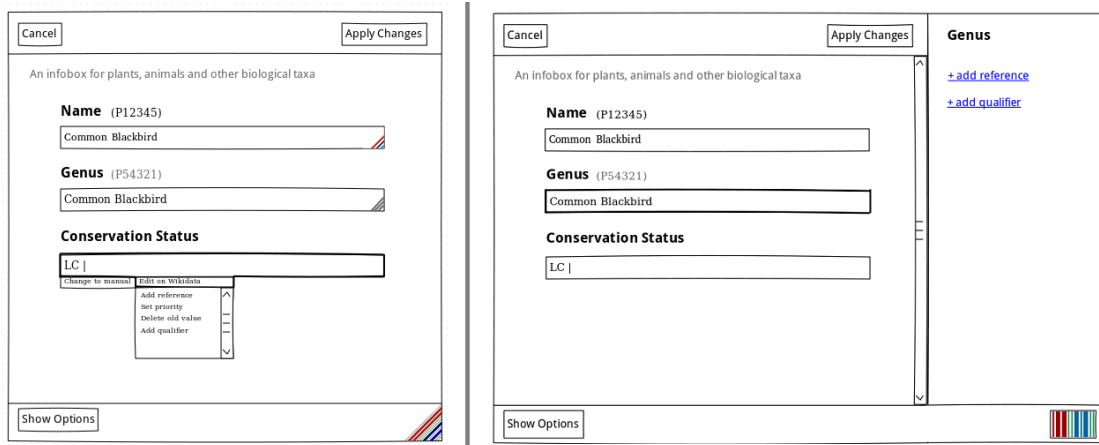

FIGURE 4.1: *Initial sketches*: on the left the drop-down version, on the right the sidebar version.

As one can see from Figure 4.1, there is a little corner symbol which is supposed to make the toggling between the two modes possible. When gray, one is in the normal Wikipedia mode and when clicked, it changes its colors to match the Wikidata design and Wikidata mode becomes activated. Furthermore, there are two tabs on the button of each field which become active when Wikidata mode is activated. The selected one at that time will be *Edit on Wikidata*. The other one says *Change to manual* and would lead back to editing manually in Wikipedia mode.

When clicking on the tab, a drop-down menu will open, which reveals several options, which are also available when editing Wikidata. One has the options to add references to the values, add qualifiers to the values, delete the entire value and thus potentially the statement, and change the rank of the value. The Wikidata corner at the bottom of the infobox changes all possible fields to get their data from Wikidata and clicking it again will revert all fields back to the manual mode.

A problem which became apparent when trying to place further required elements was the lack of space. A drop-down does not allow for a display of all current references and

qualifiers, for example, and end up being cluttered and thus illegible. Additionally. this solution would not yield enough space for tool tips and other potential information that would need to be conveyed to the user.

In the second approach, the drop-down was replaced with a sidebar, as seen on the right-hand side in Figure 4.1. This is a feature the current infobox already has, thus utilizing mechanisms familiar to the user. Currently, when clicking on the lower left button *Show Options*, a sidebar slides out to the left, offering more editing options for the user. The idea with this approach is to have a sidebar for editing Wikidata on the right, using a functionality which the editor is already comfortable with to minimize the necessary readjustment. When in Wikidata mode, by clicking the triangular stripes in the right corner of a field, the sidebar to the right slides open and reveals the tools to edit the item on Wikidata.

The functionalities found in the drop-down in the earlier draft, can now be found in the sidebar. The advantage of this design is that there is more space to display all the editable information of a statement. The sidebar can easily be retracted when the data does not need to be edited itself, but only the import of data is relevant. This is arguably also a more elegant solution, since it visually blends in easier than the drop-down.

One idea which was not further explored as part of the *Pencil* sketch was to initially have none of the aforementioned buttons and functionalities of Wikidata visible, except for one clearly defined button to activate Wikidata mode. Only when activated the buttons and functionalities become available. Every time the editor is opened, the hidden buttons would light up in a manner of a light-box, and then disappear again, leaving only the Wikidata button behind.

The problem with this concept is that the incentive to use Wikidata might not be great enough, since it is not readily available from the start and a quick switch of one field to use Wikidata would not be possible without changing into Wikidata mode entirely. Furthermore, it would not be easily visible which field uses which data, and thus lead to accidental edits, which were not intended by the user. For these reasons mainly this concept was not further pursued.

After completing the initial sketches with the sketching software, the two initial sketches are examined and evaluated. In this case, it makes sense to only implement one sketch into a functioning prototype, since creating one is a lot of work and the scope of the thesis does not allow for more than one. For reasons mentioned in the previous paragraphs, such as opening up more possibilities for edit functionalities, and the option of keeping it closed when not needed, the latter sketch is going to be realized as a prototype.

4.3 Design Process

For the next step it is important to design the specific elements and functionalities of the tool, since the sketches do not cover these with enough detail, if at all. The scope of the thesis only allows for a rather rudimentary prototype, where functionalities are not implemented in all parts, but just enough to showcase the functions and make it usable. This means that not every field gets a proper implementation of its future abilities.

The software used to create the clickable prototype is *Axure*. The prototype is modeled after the design which was previously made with *Pencil*. *Axure RP* is a wireframing, prototyping software developed by *Axure Software Solutions, Inc*. It is used to test and validate a concept in a usability test. Clickable prototypes feel real to the user, since they have no way of knowing what goes on in the background.

Using Axure makes it possible to create a clickable version of the concept fairly quickly, compared to the time it would take to implement it with code, as well as being more flexible for any necessary adjustments. The design of a prototype undergoes several minor adjustments regarding the structure and integrated features before it is tested with users in a usability test.

First, all the fields from the original infobox are copied into the mock-up, as well as all the details that are part of the elements of VisualEditor. It is important to begin with the current design as a base, in order to not deviate too far from it, since it is of utter importance to not disturb the current design too much. From there, all necessary elements can be added, or subtracted if necessary.
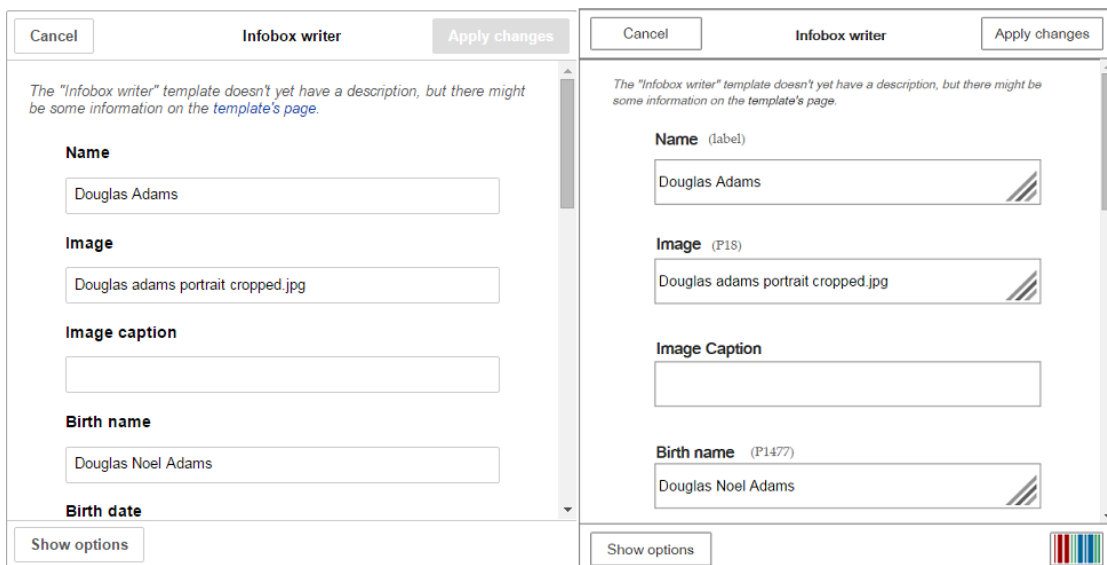

FIGURE 4.2: *Comparison*: on the left the infobox currently in use, on the right, the prototype.

28

Next, the sidebar was implemented, so whenever the corner is active, being in Wikidata mode, the sidebar opens. Since the main window usually contains more fields than fit into its predefined size, the other fields become reachable through a scrollbar. Unfortunately this scrollbar visually functions as a barrier between the selected field and the sidebar. A main goal had to be to connect the information in the sidebar visually with the selected field, so the user does not get confused about which value their currently editing.

When selecting a field, the border of the field turns blue to indicate the selection. One idea was to use this color to visually link the field and the sidebar. The title of the sidebar changes, depending on the value of the selected field. One way to link these two visually is to make the title the same color as the field border. This will help the user draw a connection between the two elements.
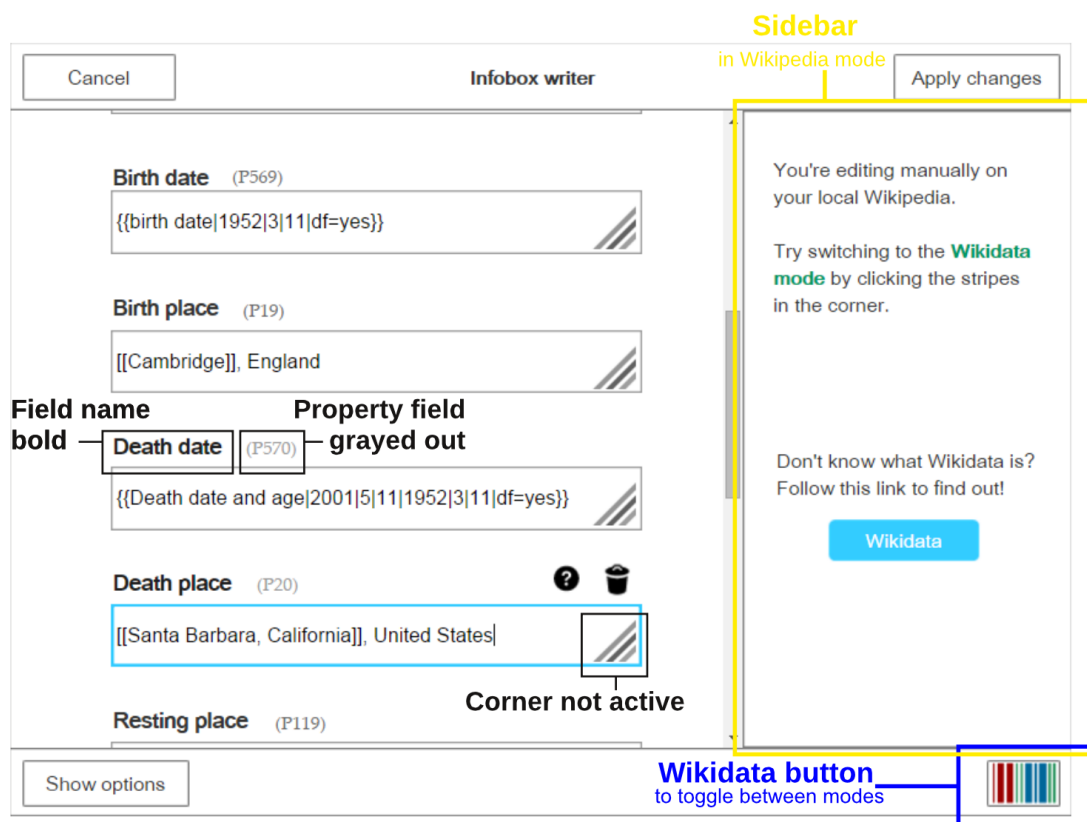

FIGURE 4.3: *Prototype*: in Wikipedia mode.

Furthermore, it is imperative to make clear to the user in which mode they are currently located. The color change in the corner alone is not enough of an indicator especially because that would not be very accessible for the colorblind. The idea is to not only have the field name written above the field, but also add in the unique identifier for the corresponding property in Wikidata.

If a field has a field name and a property field name, one for each mode, it is possible to visually toggle between the two in addition to the color change. The title corresponding

to the active mode will be displayed in bold and in the color black, whereas the inactive mode will be grayed out and displayed in a regular weight.

Since a goal of the project is not only to facilitate the use of Wikidata, but also encourage to use it, an indicator needs to be added, informing the user of the new mode, when in Wikipedia mode. When selecting a field in Wikipedia mode, the sidebar will slide open as well, displaying a short message, informing the user of Wikidata mode and how to switch to it.
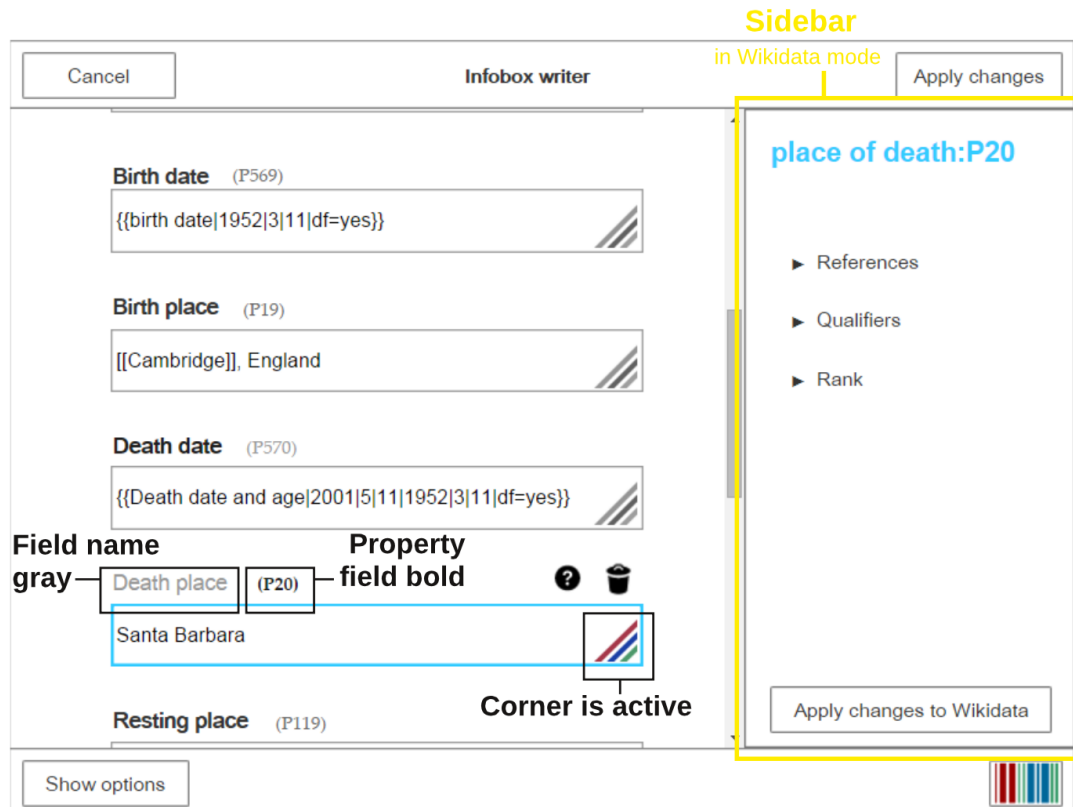

FIGURE 4.4: *Prototype*: in Wikidata mode

This message, and the sidebar in general, can be opened and removed by clicking the button in the lower right corner with the Wikidata logo. The reason for using the Wikidata logo on the button is due to the fact that it is the clearest representative of Wikidata and thus representing Wikidata mode, as well as being visually interesting so the user is inclined to click on it.

When it comes to implementing the tool tips, a few were created to support the user when they were at a loss. Looking at the information in the sidebar, it a redundancy became apparent. To test whether this information might even be sufficient, as to not clutter up the tool, the tooltips are going to be left out for the first run.

The corners for toggling between the two modes are implemented exactly like in the sketch. When clicking the corner of a field, and the user is in Wikidata mode, the sidebar changes to display the property name and the unique identifier of the property as the title in the color of the border of the field. There are three editing options for the user: Adding and editing references, adding and editing qualifiers, and changing the rank of the statement. Each of these options is collapsed. When clicking on one of them, the actual information opens, and the editing options become visible.

For instance, when the user clicks on *References* for *Place of Death* (*P20)* in Figure **??** they will see the current references already existing for that statement in Wikidata, an option to edit them, and an option to add another reference. The reference addition follows a similar process to the one on Wikidata itself. The user chooses the type of reference, in this case from a drop-down, and in another field, adds the actual reference, be it a URL or a book or another type of reference.
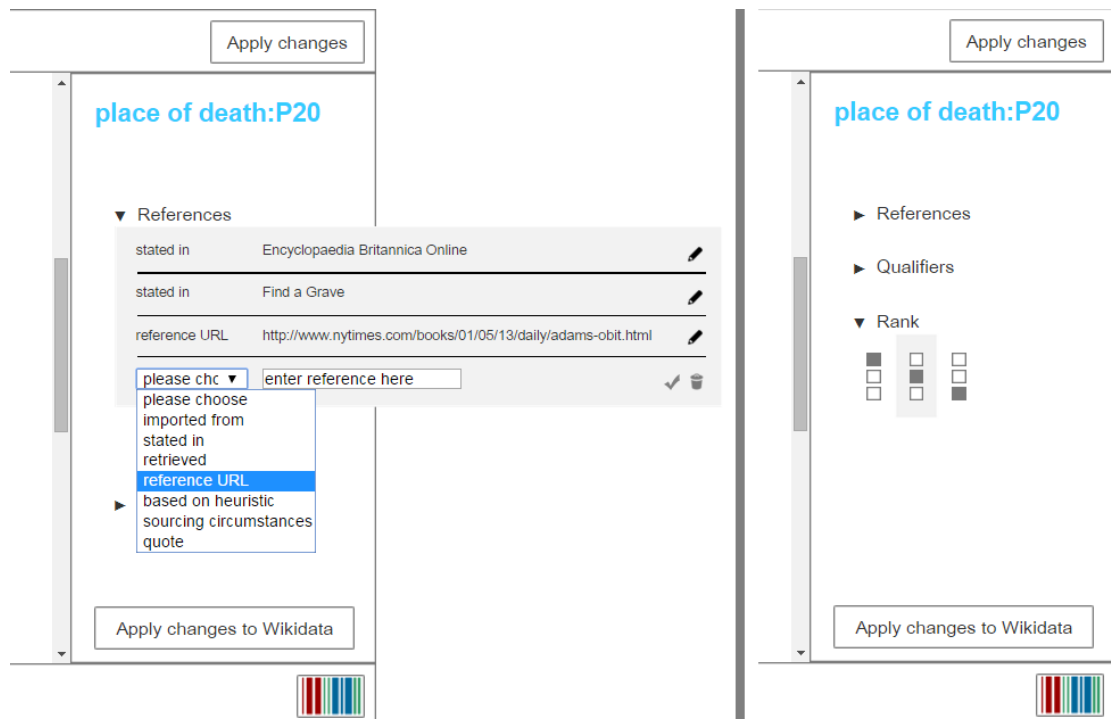

FIGURE 4.5: *Prototype*: sidebar when editing references and ranks

Qualifiers follow a similar pattern, where the ones which are already associated with the value are displayed with the option of adding new ones. The user usually picks a type of qualifier, for example a start and/or end date, and then, in an empty field, adds the actual values.

The rank of a claim in a statement can only be changed when there is more that one claim, and therefore more than one value for the field. If this is the case, the claim can be *preferred, normal*, or *deprecated*. The ranks are visualized by three squares on top of

each other, of which two are empty, and one is colored in, representing the rank of the claim amongst the other claims of the statements. It is possible to have anything from all claims ranked normal, to only one being preferred, or even all preferred except for one, since in that case, they would all be ranked normal.

The changes made to the actual item, in this case *Douglas Adams* need to be saved separately from the actual changes made to the infobox. The editing of the data on Wikidata and the integration into the infobox, and editing there represent two individual actions. Additionally, an extra step to saving the changes also works towards minimizing accidental and purposeful vandalism.
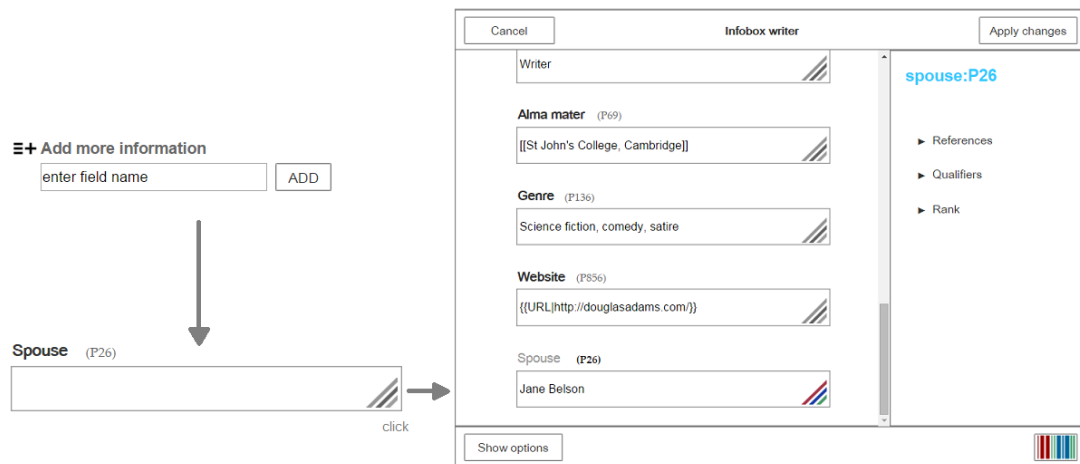

FIGURE 4.6: *Prototype*: adding another field to the infobox.

When adding a new field, one will get suggestions of other fields available to that template in a auto completing drop-down. Once the preferred field is chosen, the user is presented with an empty text box with the chosen value. This decision is primarily to prevent vandalism on Wikidata. Were the text field immediately filled with the value from Wikidata it might lead to confusion if the user is not aware of the changes made, which could easily lead to accidental vandalism. Furthermore, inserting this extra step to get to the Wikidata value also acts as a further means to inhibit intentional vandalism.

Chapter 5

Evaluation

After ensuring that all requirements have been met by the prototype, it can be tested by potential users who fit the user group. This procedure is called *usability testing* [22, p. 135]. The goal of such a test is to evaluate the usability of the product and identifying problems. The main criteria with which the success is determined is overall effectiveness, efficiency and satisfaction.

This usability test mainly tries to collect heuristic data, which is a type of analytical evaluation, compared to empirical evaluation. The latter concerns itself solely with the result of the usability test, where it does not play a role how these result came to be. Analytical evaluation concern itself with the specifics and the work-flow. If we take the usability of a pen for example. The analytical approach would be to look at what the pen is made of, the ergonomics of it, how it glides over different surfaces and so on. An empirical evaluation would only focus on how many words per minute the user can write with it and how long it lasts before being used up or broken for example.

5.1   Usability testing

5.1.1   Structure and Planning

The goal is to test a group of about 5-10 people with a varying degree of experience in editing on Wikipedia and Wikidata. Every participant is questioned separately using the same questions and tasks. The screen and mouse movements, as well as the voice was recorded with a free software called *CamStudio Recorder v2.7* for later review if necessary. After the scenario was set for them and they were informed about all formalities the test started.

First, every user was asked to describe what they see without clicking. This was useful, in order to give them some time to immerse themselves in the new environment and get a feeling for the product, before having to use it straight away. Additionally this is helpful to get a feeling for the level of comfort and experience the user has. Next, each user was asked to click around freely while restraining from editing anything. It makes sense to make this an extra step after the first one, so when the user starts clicking about, they are already familiar with everything to some extent and can make more educated decisions about where they want to go next.

Following these two steps, they were given 7 small tasks and where observed while completing them. They ranged from very easy to moderately advanced. After completing the tasks the users were asked another set of questions on their opinion of the product. Questions like *What exactly did you like, What did you not like, What was missing* and so on. These questions are helpful in determining the overall satisfaction of the user and often bring further requirements to light.

Lastly, each user was asked to fill out a short questionnaire that will help mapping them to a persona and thus evaluating if their needs have been met. These scales were similar to the one that the personas got mapped on in their development. More detailed information on the questionnaire and further materials of the usability test can be found in Appendices H - J.

5.1.2 Results

For the usability test, ten voluntary participants were chosen to cover a large range in regard to editing experience in Wikipedia, ranging from no experience to a very high experience as one can see from figure 5.1.
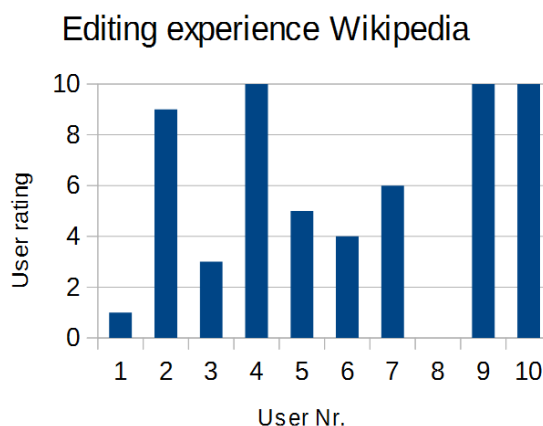

FIGURE 5.1: *Results from usability test*: Graph showing the participant's editing experience on Wikipedia.

34

Furthermore, the range of knowledge of Wikidata was also represented throughout all ranges as one can see in figure 5.2.
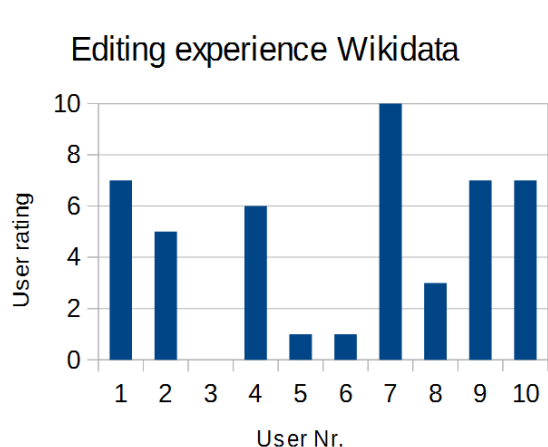

FIGURE 5.2: *Results from usability test*: Graph showing the participant's editing experience on Wikidata.

The questioning times ranged from twenty minutes to one hour, yielding various amounts of helpful information. Each participant was faced with every question, however not all questions were answered by each participant. What one can gather from the answers is that people find it easier, pointing out flaws that bother them, yet have a hard time naming the features they like. If a question was not answered at all, one can assume that the opinion of the user towards it is neutral.

5.1.3   Functional

Starting with the functional requirements, it turns out that most of negative points that users brought up were of a functional nature, since these are the easiest to name when questioned.

The issue addressed by most users, 7/10 times, was the lack of tooltips. As mentioned in section 4.3, the tooltips were purposefully left out to see if the explanation in the sidebar would be sufficient, hence seeing if the design was able to convey a clear enough picture. The initial intentions were to satisfy the principle of aesthetic and minimalist design, by not overcrowding the interface and minimizing redundancy. As it turns out, the lack of tool tips was a negative aspect to most users, since they were confused at times about the meaning of certain elements, especially the property title above the text field. Users unfamiliar with Wikidata, such as 3, 5, 6, and 8 all did not know what the p stood for since no explanation was given.

The same goes for the ranks. If a user is unfamiliar with Wikidata, they will not know what ranks are and how to properly use them. After discussing this with Lydia Pintscher,

the product manager of Wikidata, it was decided to rework ranks in the editor altogether. This, as well as all other points made in this chapter, will be further discussed in section 5.2.

Another feature the users missed regarding to images was a thumbnail preview of the image selected, so the user can be sure, that the correct image was chosen for the infobox before committing any changes. This was not really a criticism of the implementation, but rather an new requirement that only became apparent to the users when using the tool.

The by far most misunderstood design element was the Wikidata button in the lower right corner. 7/10 users expected to be redirected to the Wikidata web page or the respective item on Wikidata when clicking the button. In chapter 4.3 it was established that it represents Wikidata the best, since it is the actual logo and could easily stand for the Wikidata mode but the association with Wikidata directly seems to be too strong. It did not suggest to them, that it might be used to switch between the two modes in the editor. This can partially be attributed to standards, since users generally expect to be taken to the main page of the project or site behind the logo, thus it was recognized as a link rather than a button with a functionality in the infobox.

A second main issue with standards became clear, which was a misinterpretation of the corner symbol on the text field. When in gray mode 3/10 people thought that it was a text field size adjusting element. This is a clear indicator of the failed fulfillment of the conformity with the expectations of the user. This is clearly something that needs to be addressed in the next iteration, since this functionality is a key element of the tool and not being recognized as such can be very harmful to its usability.

Another issue were the colored words in the sidebar. 3/10 users recognized them as links and tried to click them in order to switch to Wikidata mode. This is not a major issue, but would still overall have a negative impact on the user experience.

Furthermore, 4/10 users mentioned that they were not certain enough —through the feedback of the tool—if their changes were saved at all, and if they would be informed if they were to exit the infobox with unsaved edits. It became apparent, that the visibility of the status was not being communicated sufficiently to the user.

On the positive side, 4/10 users clearly stated that they found the toggling between the two modes was being communicated very clearly. The change in the colors and the boldness of the text as well as the change in the sidebar helped them with placing themselves locality wise i.e. being aware of which mode they were in.

5.1.4   Non-Functional

Unlike the negative findings, most positive ones were of the non-functional kind.

Starting with the negative comments, the main issue that stood out was a navigational one. 3/10 users put forth difficulties "trying to get back" to the Wikipedia mode or when trying to close the sidebar. This goes hand in hand with the misinterpretation of the Wikidata button [a] since that was the navigational link they were missing. This is a prime example of a lack of *user control freedom* where the user feels trapped in some way by the product and does not have the feeling that they are fully in control of the navigation.

3/10 users voiced their concerns regarding the impact of their edits. To be precise, when editing Wikidata, they were uncertain if the edit was also happening on Wikidata directly or just locally on Wikipedia. This is a result of a lack of feedback to the user, thus the *visibility* of the state of the system is not communicated clearly enough to the user.

Moving on to the positive feedback, 2/10 users pointed out that they find the vandalism barrier was high enough to be successful, while feeling that the measures that were taken were not interfering with their editing work.

Furthermore, 7/10 users commented that they appreciated the easy access to the items on Wikidata through the automation, hence providing a greater level of comfort when making edits which will significantly facilitate their work. Furthermore, they pointed out that the editing power, which the tool gives, makes it very comfortable to edit. 5/7 of these users were the ones who gave themselves the highest rating out of the 10 participants on the questionnaire for the last two questions regarding Wikidata.

Many of the positive comments were directed at the immediate design decisions. 3/10 users found the colors, the graphical elements and the general structure visually pleasing or just generally appreciated the design details. This may not seem as important as the previous points, however creating a visually pleasing product contributes to a successful product just as much as it minimizes work for the user. [2, p.92]

Another comment was that 2/10 participants pointed out that they appreciated the quick input possibility of references, which is a step that should be entirely barrier-free, since referencing should be encouraged where possible.

5.2   Evaluation

Focusing on the information gathered from the design process, following measures need to be taken to improve the prototype:

The lack of tooltips has become very clear in multiple places during the usability tests. Mainly all Wikidata related specifics need to be explained separately, like the property identifiers, qualifiers and ranks. Having information on the usage in the sidebar was not sufficient, and thus needs revision.

The Ranks should be removed the way they are implemented currently, since it was decided not to allow for deprecation and it is difficult to convey the concept of them with little effort. Instead the idea is to have a check box for every value, to verify if there is supposed to be a preferred one. If nothing is checked, the default, which is ranked normal, is active. In the template for certain properties the check box would not be available at all, as for example for the property *children* since there can not be a preferred one. The template editors should also have the possibility to have certain fields in their infobox demand at least one preferred value, for example the property *mayor* in a city infobox or the *population number*.

The next point is about the representation of values. It needs to be possible to represent multiple values per field like it is possible in Wikidata. For the current prototype the aspect of editing several values for one field has not really been considered, since this is not often the case and it possibly is the most complicated part to implement.

The idea is to have all values as separate entities in the field, and each of them will be clickable to then edit them in the sidebar. There is an option to set the rank with each of them individually, but this would mean, when changing ranks one would have to go through all of them in the worst case, and find the one that is currently preferred and set it back, in order to set another one to preferred. It could not just reset the other value automatically, since it is possible to have multiple values at this rank.

Of course the preferred value could be highlighted visually in the field so the search would not take up any time. The other option would be, when clicking into the field, but not on a specific value, the sidebar would open all values side by side and then let the user set a preferred value there, when having an oversight over all the values. This would of course require an extra step and the user might not find it since he would have to click in between the values to get to this value display.

Moreover, the Wikidata button should be reworked and the functionality of it should be made clearer, meaning that it needs to showcase more that it makes the sidebar slide out and that it toggles between the two modes. The button itself as a tool should definitely

stay. A possibility could be to present users with a handful of options regarding the style of the button, to find out which one conveys the message best.

The corners were misinterpreted by 33 percent of the users as text field resizers. Their existence makes sense but not the design choice. The corner needs to be either turned around, or converted to vertical lines or be reworked completely. The two field labels could be reworked as tabs to allow for an even more visual toggling.

Another idea that would be great to try in the next iteration would be to make a new field according to which mode it was in, when it was created. When the editor is in Wikidata mode, chances are that when he creates the new field he would like the value already in place, unlike the editor in Wikipedia mode, who most likely does not want to make use of Wikidata.

It would probably also be important to label the modes somewhere, make it even more clear where the user is located and thus disperse the confusion about locality.

After making aforementioned changes, a new set of users would be invited to test the new prototype. The results would then be evaluated, just like it was done earlier in this thesis, and the necessary changes would be made. At this point it would be up to the team to decide, whether the product needs another round of iteration or whether the users were sufficiently satisfied, that the products could be rolled out.

Chapter 6

Conclusion and Outlook

6.1 Critical Analysis

The end result of this thesis should definitely not be seen as an end result to the whole project. The scope of the thesis did not allow for a single iteration, although it would be recommended to at least have one, or more if possible. Without re-iterating through the process after the first evaluation, there is no chance to improve the product with all the knowledge gained from the usability test. As mentioned in previous chapters this is crucial for developing user-centered.

6.2 Future Work

After the conclusion of this thesis, the project will undergo further iterations. This will include updated revisions of the prototype which incorporate the results obtained by the usability test. This updated prototype will then undergo further usability testing, to optimize the product in order to meet the previously established requirements and reach an improved effectiveness, efficiency and satisfaction for the users as discussed in section 2.4.

This process is repeated until a satisfactory level is reached on all three aforementioned points which determine the usability of the product. When this point is reached, the product will be implemented and offered as a beta feature to the English and German Wikipedias at first, followed by other languages.

At this point, there is much criticism and rejection to be expected, just as it was the case when Wikimedia Commons first launched. Nonetheless, it can be assumed that the

positive prospect the tool holds will be able to convince most community members of its potential and advantages.

Since this tool was developed for an agile environment such as Wikipedia with an active community, its development will never permanently conclude. It is to be expected that adjustments will have to be made continuously throughout the lifetime of the project, especially when adapting it to fit the needs of further Wikimedia and third-party projects.

6.2.1   Usefulness in other Wikimedia Projects

Wikimedia has several projects that could benefit from the work done in this thesis, since Wikipedia is not the only project that has a type of infobox on its page.

An example of a Wikimedia project that would benefit from the previous work is *Wikivoyage*. This project is an online travel guide and like the other Wikimedia projects, free and editable by anyone. Wikivoyage uses infoboxes on their articles to showcase the most important facts about its topic, much like on Wikipedia.

Moreover, *Wikispecies* could make use of this research. It is a free online directory on species, aiming to create a complete catalog for anyone to use. As a directory, there is no continuous text used for the articles, but only information in a structured way, which makes it perfect for use of data from Wikidata.

The last example is *Wikinews*, which is an online news source providing free content, just like its sister projects. Most news articles include an infobox which may have additional information on the topic, or headlines of other stories about the same topic. The supplementary information on the topic could be drawn from Wikidata, and hence profit from being editable from the client directly.

6.3   Conclusion

The implementation of a user interface within other Wikimedia projects to edit Wikidata has revealed itself to be a subject with many complex facets. Due to the involvement of the community and having to design onto an already running and established system, the requirements were rather strict, thus creating a challenging design process. Nevertheless, this thesis has created a solid basis which, after further iterations, will be implemented into various Wiki projects in the future.

Implementing the solution this thesis has brought forth will provide a remedy to the current situation where the data across the different Wiki projects and all the language

Wikipedias is incoherent and partially inaccurate. With the hurdle removed, of having to leave the local Wiki to make edits, the users can now comfortably edit on the spot and do not have to revert to using local data for comfort reasons.

The major advantage of a localized system is that not every community needs to know about everything. Smaller language Wikipedias can easily use correct data maintained by editors or communities concerned with a specific topic. This results in more time each community and editor can dedicate to the interests and subjects relevant to them.

Not only would the implementation of the concept of the thesis result in more coherent and accurate data for all Wikimedia projects, but it would also increase the quality and quantity of data on Wikidata. Editors from other Wiki projects, especially Wikipedia, would be adding new data directly to Wikidata instead of storing it locally. Furthermore, it would then automatically be maintained by these communities resulting in a —one could say—symbiotic relationship.

Overall, this project could be the missing puzzle piece for Wikidata, to take its usage to the next level and thus helping it rise to its full potential and improve the data quality across all Wikimedia projects.

List of Figures

Bibliography

[1] 1.0, CC0: *Public Domain Dedication.* https://creativecommons.org/publicdomain/zero/1.0/, 2014

[2] ALAN COOPER, Robert R.: *About Face 2.0 - The essentials of interaction design.* Wiley Publishing Inc., 2003

[3] BUTLER, Christopher: *How to Tell the User's Story.* https://www.newfangled.com/how-to-tell-the-users-story/, . – [Accessed 4. December 2015]

[4] CALABRIA, Tina: *An introduction to personas and how to create them.* http://www.steptwo.com.au/papers/kmc_personas/, . – [Accessed 18. December 2015]

[5] COMMONS, Wikimedia: *Commons:Welcome.* https://commons.wikimedia.org/wiki/Commons:Welcome, . – [Accessed 06. November 2015]

[6] COMMUNITY, English W.: *Wikipedia:Wikidata.* https://en.wikipedia.org/wiki/Wikipedia:Wikidata, . – [Accessed 12. January 2016]

[7] COMMUNITY, Wikidata: *Wikidata:Client editing input.* https://www.wikidata.org/wiki/Wikidata:Client_editing_input, 2015. – [Accessed 30. October 2015]

[8] COMMUNITY, Wikipedia: *Help:Infobox.* https://en.wikipedia.org/wiki/Help:Infobox,

[9] COMMUNITY, Wikipedia: *Why Wikipedia is not so great.* https://en.wikipedia.org/wiki/Wikipedia:Why_Wikipedia_is_not_so_great, . – [Accessed 13. December 2015]

[10] COMMUNITY, Wikipedia: *Wikipedia talk:Moving files to the Commons.* https://en.wikipedia.org/w/index.php?title=Wikipedia_talk:Moving_files_to_the_Commons&oldid=29967943,

[11] CONSORTIUM, DSDM: *The DSDM Agile Project Framework (2014 Onwards - Chapter 10).* https://www.dsdm.org/content/moscow-prioritisation, 2014. – [Accessed 14. February 2016]

[12] EN, Wikipedia: *Help:Interlanguage links.* https://en.wikipedia.org/wiki/Help:Interlanguage_links, . – [Accessed 05. January 2016]

[13] EN, Wikipedia: *Template:Infobox person.* https://en.wikipedia.org/wiki/Template:Infobox_person, . – [Accessed 06. January 2016]

[14] EN, Wikipedia: *Wikipedia:User access levels.* `https://en.wikipedia.org/wiki/ Wikipedia:User_access_levels`, . – [Accessed 02. February 2016]

[15] EN, Wikipedie: *Wikipedia:About.* `https://en.wikipedia.org/wiki/Wikipedia:About`, . – [Accessed 01. November 2015]

[16] EN, Wikipedie: *Wikipedia:VisualEditor.* `https://en.wikipedia.org/wiki/Wikipedia: VisualEditor`, . – [Accessed 28. November 2015]

[17] EN, Wikipedie: *Wikipedia:Wikipedians.* `https://en.wikipedia.org/wiki/Wikipedia: Wikipedians`, . – [Accessed 05. December 2015]

[18] Florian Sarodnick, Henning B.: *Methoden der Usability Evaluation - Wissenschaftliche Grundlagen und praktische Anwendung.* Hans Huber, 2011

[19] Foundation, Wikimedia: *Frequently asked questions.* `https://wikimediafoundation. org/wiki/FAQ/en`, . – [Accessed 03. November 2015]

[20] Gothelf, Jeff: *Using Poto-Personas for Executive Alignment.* `https://uxmag.com/ articles/using-proto-personas-for-executive-alignment`, May 2012. – UX-Mag

[21] Kolbe, Andreas: *Whither Wikidata?* `https://en.wikipedia.org/wiki/Wikipedia: Wikipedia_Signpost/2015-12-02/Op-ed`, December 2015. – The Signpost

[22] Krug, Steve: *Don't make me think - A Common Sense Approach to Web Usability.* New Riders, 2006

[23] Mayhew, Deborah J.: *The Usability Engineering Lifecycle - a practitioner's handbook for user interface design.* Morgan Kaufmann Publishers, Inc., 1999

[24] Nielsen, Jakob: *10 Usability Heuristics for User Interface Design.* `https://www.nngroup. com/articles/ten-usability-heuristics/`, 1995. – [Accessed 19. Februar 2016]

[25] Perez, Sarah: *Wikipedia's Next Big Thing: Wikidata, A Machine-Readable, User-Editable Database Funded By Google, Paul Allen And Others.* `http://tcrn.ch/H0aO9U`, . – [Accessed 06. February 2016]

[26] Pintscher, Lydia: *Wikidata: Knowledge from different points of view.* `https:// en.wikipedia.org/wiki/Wikipedia:Wikipedia_Signpost/2015-12-09/Op-ed`, December 2015. – The Signpost

[27] Pressemitteilung, Wikimedia G.: *DATA REVOLUTION FOR WIKIPEDIA.* `https: //www.wikimedia.de/wiki/Pressemitteilungen/PM_3_12_Wikidata_EN`, March 2012. – [Accessed 06. February 2016]

[28] Standardization, International O.: *Ergonomic requirements for office work with visual display terminals (VDTs) — Part 11: Guidance on usability.* `https://www.iso.org/obp/ ui/#iso:std:iso:9241:-11:ed-1:v1:en`, 1998. – [Accessed 14. December 2015]

[29] STANDARDIZATION, International O.: *Ergonomics of human-system interaction — Part 110: Dialogue principles.* `https://www.iso.org/obp/ui/#iso:std:iso:9241:-110:ed-1:v1:en`, 2008. – [Accessed 1. December 2015]

[30] STANDARDIZATION, International O.: *Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems.* `https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-1:v1:en`, 2010. – [Accessed 14. December 2015]

[31] USABILITY.GOV: *Personas.* `http://www.usability.gov/how-to-and-tools/methods/personas.html`, . – [Accessed 17. December 2015]

[32] USABILITY.GOV: *Scenarios.* `http://www.usability.gov/how-to-and-tools/methods/scenarios.html`, . – [Accessed 4. December 2015]

[33] VOS, Dick de: *By Dick de Vos (Own work) [CC BY-SA 4.0 (http://creativecommons.org/licenses/by-sa/4.0)], via Wikimedia Commons.* `https://commons.wikimedia.org/wiki/File%3ADick_de_Vos.JPG`, 2011. – [Accessed 1. November 2015]

[34] WIKIDATA: *Help:Bots.* `https://www.wikidata.org/wiki/Help:Bots`, . – [Accessed 13. October 2015]

[35] WIKIDATA: *Help:Items.* `https://www.wikidata.org/wiki/Help:Items`, . – [Accessed 13. October 2015]

[36] WIKIDATA: *Help:Properties.* `https://www.wikidata.org/wiki/Help:Properties`, . – [Accessed 13. October 2015]

[37] WIKIDATA: *Help:Qualifiers.* `https://www.wikidata.org/wiki/Help:Qualifiers`, . – [Accessed 13. October 2015]

[38] WIKIDATA: *Help:Statements.* `https://www.wikidata.org/wiki/Help:Statements`, . – [Accessed 13. October 2015]

[39] WIKIDATA: *Wikidata Main Page.* `https://www.wikidata.org/wiki/Wikidata:Main_Page`, . – [Accessed 25. February 2016]

[40] WIKIDATA: *Wikidata:Introduction.* `https://www.wikidata.org/wiki/Wikidata:Introduction`, . – [Accessed 18. October 2015]

[41] WIKIMEDIA: *Grafana - Wikidata Entity Usage.* `https://grafana.wikimedia.org/dashboard/db/wikidata-entity-usage-project`, . – [Accessed 01. February 2016]

[42] WIKIMEDIA: *Wikipedia Statistics - Active Wikipedians.* `https://stats.wikimedia.org/EN/TablesWikipediansEditsGt5.htm`, . – [Last accessed 29. February 2016]

[43] WIKIMEDIA: *Wikipedia Statistics - Edit activity levels of registered users.* `https://stats.wikimedia.org/EN/TablesWikipediaZZ.htm#editor_activity_levels`, . – [Last accessed 29. February 2016]

[44] WIKIMEDIA: *Not my wiki.* `https://meta.wikimedia.org/wiki/Not_my_wiki`, June 2014. – [Accessed 02. February 2016]

Appendix A

Persona: Miriam Weiland

Demography:

- Female

- 36 years old

- Lives in Heidelberg, Germany

- Married

- Studied veterinary medicine Uni Heidelberg

- Lives in a single apartment in quite neighborhood

- Works as a veterinarian at Uniklinik Heidelberg

- Has two dogs

- Apartment in city center

Behaviors:

- Calm personality

- Very assertive

- Very sociable

- Pragmatic/ goal oriented

- Helper personality

- Spends a lot of time in forums answering users questions

Needs and Goals:

- Protecting animals

- Spread awareness

- Likes the feeling to be needed and help

- Would like more regular working hours

- Worries that her marriage might fall apart

- Scared of failing at her job and thus hurting someone

Miriam is 36 year old veterinarian working at the University clinic of Heidelberg. She often has to work long shifts and sometimes also night shifts. Often when she comes home her husband is already asleep and so their marriage has not been very stable for a while. She's thinking of changing to a regular veterinarian's office so she can have normal working hours. She lives for her job because helping animals is what drives her. When she sees a suffering animal, that needs help, she is unable to resist. She tries to extend her knowledge of animals, their illnesses and their potential treatments to the outside world because she wants that as many animals as possible find help. That's the reason why she spends hours in forums, when she comes home from work, answering people's questions about their animals.

A couple of years ago she discovered Wikipedia as a tool to also spread knowledge. When she started editing she found that many articles on animal illnesses were incomplete or entirely missing. Over the years she has written and edited 20+ articles hoping that someone will make use out of them. Most of her articles are rather scientific and thus often require an infobox. Miriam knows about Wikidata but has no idea how to integrate the information and no time to figure it out. Right now she always types everything by hand but she would appreciate the help of not having to do so by just integrating Wikidata's item. She also already encountered the situation where there was an English version of her article but she did not know how to connect both of them. She often worries that people don't take care properly of their pets. She has seen many situations in her work where mistreat them without even them knowing it.

She sometimes thinks about having kids, but the way her marriage is going she does not think that would be a good idea. She feels bad for not being home often and letting her marriage slip but she is incapable of giving up on her animal care. She is determined to make a difference, even when it is just one article at a time. She loves her two dogs, Lilo and Stitch. She tries to spend as much time with them as possible taking them for extensive walks as often as she can and taking them on her run every morning.

Expectations and desires:

- Make editing easier

- Less information to update

- Save her time

- Easily understandable

- Would appreciate an auto-generation button for the infobox

User scenario:

1. After Miriam comes home from work she sits down to answer questions that people might have asked her in one of the many forums that she's active in. In response to a specific issue she wants to link the relevant Wikipedia article. She notices a small mistake in the infobox though and quickly wants to change it, before sending the link. She logs in and opens the VisualEditor as usual. She sees that the value used in there is still typed in by hand. She much prefers the infobox to use Wikidata so she goes ahead and switches the box to "use Wikidata item for this value" and when she saves the correct value for that property it's displayed in the infobox.

2. She sits down to do some research on a disease she encountered today at the vet's office. She finds some information in veterinary journals but sadly no Wikipedia article. She knows that people would never read those journals to get informed about a disease so she decides to write an article. She writes the text first which she derives from multiple of the scientific papers she has read on this issue. It always takes so much time to write the articles accurately and not get any of the facts wrong so all she really wants is to have to do as little as possible. The new function of auto-generating an infobox is exactly what she wants. She just has to open the visual editor and click on create infobox from Wikidata item, type in the Item ID and the infobox template she wants to use and everything is generated for her. She then goes through all the information to make sure it's correct before saving.

3. There is a short article on the English Wikipedia, so she wants to make sure later, to link those two together. She used to have to do this manually, but because she used the Item in the infobox generation it automatically suggested to her to add this article to the item since it realized that there is no German article linked to this item yet. She clicks yes and is relieved that yet another task is taken of her hands. She gets to go to bed earlier than expected which makes her very happy.

Appendix B

Persona: Sam Doughty

Demography:

- Male

- 22 years old

- Lives in Leeds, UK

- Studies physics

- Parents support him financially

- Freelances a bit on the side doing some minor programming work

- Loves to bike

Behaviors:

- Does not develop emotional attachments easily

- Relaxed/easy going

- Very sociable

- Lives into the day

- Avoids most confrontation

- Night owl

- Hangs out a lot with friends and gf

- Never hangs on to something for a long period of time

Needs and Goals:

- Explore the world

- Do something new all the time

- Fears that he will never find something that fascinates him enough

- Would like to be able to hold on to one thing for long

- Worries that he won't finish his studies

- Wants more challenges in life

Sam is a 22 year old student who lived in Leeds all his life and is now studying physics at the local university. He loves to take his bike everywhere and since the city is rather small that's usually not a problem. He has always been very uncertain about most things in life and has changed his major twice. He has started playing 4 different music instruments already and always quit after a couple of months. He's in university listening to his professor speak about *spaghettification* in regards to black holes and for once finds it easy to listen to the lecture. After he gets home he checks Wikipedia to read up on it more. He finds that there is no section on the black hole article yet so he decided to edit it. He mistypes his password a couple of times because he has not logged in in such a long time. He likes editing but only from time to time and when there's a reason. He can't be bothered with always keeping everything up to date and having to worry about whether the information is still accurate but when he has something to say he enjoys writing it down for everyone to read. He's meeting his girlfriend in 30min at the cinema to watch the new bond movie. After they get home he reads up on the article of the movie and notices that it's missing some crucial information which he then quickly adds. He sometimes worries that he'll never find anything that he loves so much, that he'll stick with it for longer. At least he has been editing on Wikipedia sporadically for a while now.

Expectations and desires:

- Intuitive enough to understand after not having used for a long time

- What is Wikidata

- Wants not to think just do

User scenario:

1. In between two appointments Sam wants to quickly edit an infobox of an article. He only has very little time of which he spends 5 minutes remembering his login details. When he goes to edit the information in the visual editor he quickly sees that the value is not derived from Wikidata. He changes the settings to use the information from Wikidata and the information in the infobox is corrected automatically.

2. When he goes to change another mistake he finds that the property's value is already taken from Wikidata but the value is still wrong. From the drop-down he selects that it's a wrong value. He's asked if the value is inherently wrong or was ever right at one point. He picks the first and is asked to type in the correct value. Suggestions are made for him. He picks one but unfortunately does not have a reference at hand. Before he selects 'ok' he is asked whether he is sure that he wants to remove the old value from that property. He selects 'ok' and the new value appears in the box.

3.

Appendix C

Persona: Sven Kueppner

Demography:

- Male

- 28 years old

- Lives in Mannheim, Germany

- Single for 2 years

- Studied economics at Uni Mannheim

- Lives in a single apartment in quiet neighborhood

- Works at a white collar job in a non-leading position

Behaviors:

- Gets into heated discussions a lot

- Has problems handing of control

- Thinks his opinion is being overheard a lot of the time

- Very suspicious of people and their Behaviors

- Avid *twitter* and *facebook* user

- Likes to comment on most things to display his opinion

Needs and Goals:

- Receiving the respect he feels he deserves

- Wants more control and responsibility

- Wants to be admired

- Wants to have something to show for

- worries that he won't "make it" in life

- Scared of being compared to other people

https://www.wikidata.org/wiki/User:Sicherlich

"If you find (how ever) your way from the Wikipedia-Artikel Warszawa to Q270 - then what? If I would like to use the coordinates then what am I supposed to do? clicking on it? - no. edit? - no. Well so what? how? ? - the free encyclopedia that anyone can edit...."

Sven is a 28 year old male who works as an accountant in a big company. He is the younger brother of two and always strives to be as successful as his older brother. He does not like the level of responsibility he has at work and always tries to impress everyone with his knowledge. Usually his co-workers ignore what he has to say or don't take it too seriously. That really upsets Sven but he feels like he's powerless against it. He feels like he's really good at his job and deserves a promotion. He wishes that his co-workers would acknowledge his abilities and thus would look up to him more. When his brother was his age he was already head of a department of his company. That really bothers him. When he gets home from work he tries to meet a friend for drinks but he's busy. He feels like lately his friends have less and less time for him because they all seem to be so occupied with their jobs and their relationships. Since his relationship ended two years ago, he has had a hart time meeting the right person. This only adds up to his frustration. He knows that it is not healthy but he does not know what to do about it. Even in his editing community he is not very respected and he's been editing for over 6 years. He likes to participate in discussion because he feels like that way he can make a change, but often the reaction to his comments is not very positive. He hates it when people mess with his articles. He does not like sharing the editing powers although he knows that Wikipedia is a community contribution but he feels like it's almost a personal violation when his words are being edited. He usually maintains the English and German Wikipedia articles and is often annoyed that he has to retype the change twice when things change. He is very skeptical about the infobox information coming from Wikidata but on the other hand it would really relieve him from a lot of typing.

Expectations and desires:

- Wants to keep a say in what will be written there

- Does not want it to be easier for other people to edit his article

- Straight forward Wikidata integration

- Does not want to be made feel stupid

- No transfer to Wikidata itself!

User scenario:

1. When Sven comes home from work he goes to check on his ongoing discussions with other editors. He sees that someone edited something in one of his articles again. There has been an ongoing discussion for a while for a specific value in the infobox on which no one can agree. Sven decides that it's time to settle this now and choses to finally implement Wikidata for his infoboxes. He opens VisualEditor and uses the convert to get information from Wikidata button. All the properties that are available are changed to receive their information from there. Fortunately for Sven the value which was disputed, is the same as his in the Wikidata item, so it is now changed back to his desired value.

2. Later that evening Sven realizes that an article he follows does not have an infobox so he decides to add one. He opens the VisualEditor and decides to just let it automatically generate it, although he does not fully trust that yet but it's just one button click away and he can still edit all the information afterwards if necessary.

3. He is surprised at how easily the changes are possible without requiring any knowledge about Wikidata. He decides to go through some more infoboxes to see if Wikidata is being used yet. He finds one where most values are still manually written in. He sets all possible boxes to 'get from Wikidata'. Before hitting the 'save' button he can see how it would look with the values to make sure everything is correct. He really likes that feature so he does not have to commit to something that might be wrong.

Appendix D

Persona: Peter Wellington

Demography:

- Male

- 54 years old

- Lives in San Francisco, USA

- Married with kids

- Software developer

Behaviors:

- Particular about the way he likes things

- Does not like to take orders

- Used to being in a leading position

- Appreciates quiet and calm

Needs and Goals:

- Spreading free knowledge

- Slow and steady wins the race

- Fears that people oversee the obvious because everyone is rushing everywhere

- Prefers emails to instant messaging

- Worries that there's too much constant input overwhelming our senses

- Prefers everything to be neat and sorted

Peter is a mid-fifties software engineer who has been working in the industry for so long, not much can faze him. He does not really appreciate hectic surroundings and prefers it to be quiet so he can work efficiently. He's in control of the software department at his company and runs everything steadily. He believes that slow and steady wins the race although he is not opposed to improvements. His day is usually pre-structured and he avidly avoids makes any chances to it. It does not necessarily bother him. He is a very calm and stable person and there is not much that can shake him up. He has been editing Wikipedia articles for 6+ years but never really got into it. He never liked how everyone could just come in and change everything. He often finds that the guidelines for the English Wikipedia are not strict enough or at least not enforced rigorously enough. In the last couple of years he discovered template creation for himself. Usually no one tampers with his templates and everyone uses them when making an article which makes him feel good. When people want changes it'll be discussed before they are put in place so it's not as much out of control like an article would be. The only thing that worries him when it comes to being able Wikidata from the VisualEditor is that people will want to change loads of templates significantly. Other than that he is not opposed to the idea of it.

Appendix E

Persona: Tanja Brinks

Demography:

- Female

- 33 years old

- Lives in Berlin, Germany

- Open data journalist

Behaviors:

- Heart involved with Wikidata and Wikimedia projects in general

- Fights "the good fight"

- Very tolerant

- Rather shy

- Has problems speaking up

- Goes unnoticed at times

Needs and Goals:

- Spreading free knowledge

- Creating loads of freely accessible data

- Keep Wikidata vandalism free

- Prefers all online contact over real life situations

- Wants Wikidata to be used more avidly

- Worries that it does not get the attention it deserves

- Believes that Wikidata has the potential to change the future

Tanja is a 33 year old journalist who mostly sticks to herself and only goes out when necessary. The topic she is specialized in is Big Data and open data. She believes that this type of data is the future and it needs to be expanded and made machine readable as soon as possible. This is the reason why she supports Wikidata so strongly. She thinks this might be the most important collection of machine readable data ever. She vigorously edits and patrols the data and always tries to encourage people to use it. She is not always very successful at that because she is a very timid person and has trouble talking to people and often is not taken too seriously. She always strives to please everyone and has trouble confronting people about things. That is especially a problem in hes business because you need to be quite persistent and memorable to be successful. That's why she tries to stick to the more online approach. She loves the idea of Wikidata being used in Wikipedia more and integrating the editing functionality into the VisualEditor. The only worry that she has is that it might facilitate vandalism of the data but she is definitely willing to give it a try if that means that the data will get more use and the knowledge of it's existence will be spread further.

Appendix F

Persona: Jamie Morgenstern

Demography:

- Female
- 20 years old
- Lives in Gotenburg, Sweden
- Art student

Behaviors:

- Unproblematic personality
- Usually liked by everyone
- Enjoys people's company
- Likes to try new stuff every now and again
- Active on all social media channels

Needs and Goals:

- Being occupied
- Making friends
- Belonging to a part of a group
- Likes the feeling of being included

- Fears that she will end alone one day

- Is annoyed by trolls and purposefully evil people

- Tends to find a person to leech on to

Jamie is a 20 year old student trying to figure out what she wants in life. She's studying art right now but used to study archeology and before that architecture. Once she finds something she loves she usually sticks to it though but it does not always work on the first try. She is constantly scared to miss out on something and thus tires out many things before settling. She just started editing on Wikipedia and so far really enjoys it. She likes the community and appreciates all the help she is getting. Also she has the feeling that what she's doing actually has a impact which is important to her. She has been to a women's open editing session in her hometown already where she met loads of nice women. She is not the most tech-savy person but quite clever so she usually figures out how something works fairly fast. On the other hand she is so used to all this user-friendly tech nowadays that it's hard to tell whether she's just generation internet or if she actually has a good understanding of technology.

Appendix G

Persona: Max Senkmann

Demography:

- Male

- 17 years old

- Lives in Pittsburgh, USA

- Senior in high school

- Lives with his grandma and mother

- Likes to ride his long board

Behaviors:

- Often bored

- Acts very childish at times

- Tests out limits

- Low self-esteem

- Typically shies away from responsibilities

Needs and Goals:

- Seeks attention

- Does not like to be alone

- Wants to be respected

- Tries to impressing people

- Thinks asserting dominance will make him feel more powerful

- Fears that he is no one

- Avoids being left alone with his thoughts

Max is a bright kid but had not the best upbringing and was not taught how to deal with all his feelings. All he really seeks is a group that he can belong to. He thinks he constantly needs to impress people for them to like him, otherwise they'll just leave. He definitely has abandonment issues which are unresolved. When he has no outlet and feels alone he tried to trigger any human reactions or interactions possible. For him trolling is an easy way to establish a connection. He recently discovered that vandalizing articles on Wikipedia can be fun and triggers peoples reactions easily, and it all occurs anonymously. He knows he's not gaining anyones respect with that but at least it's some acknowledgment that he exists. He does not really know where his life is leading. He is in his last year of high school now and has no plans for later. High school is quite easy for him although he could have much better grades if he would take it all more seriously and would not mess about with his friends so much.
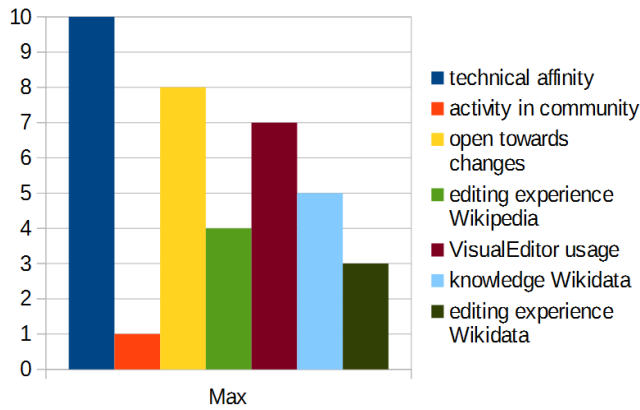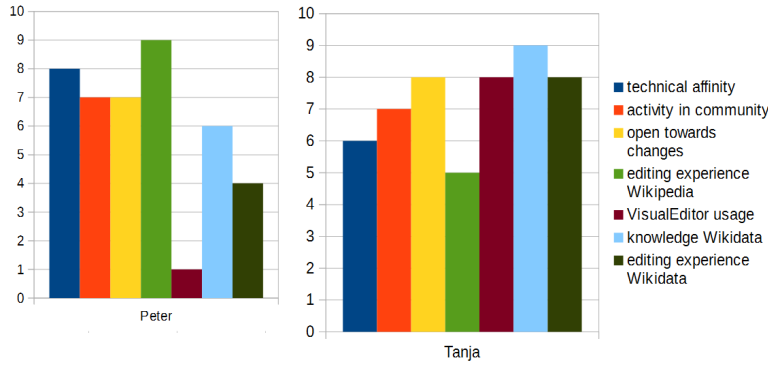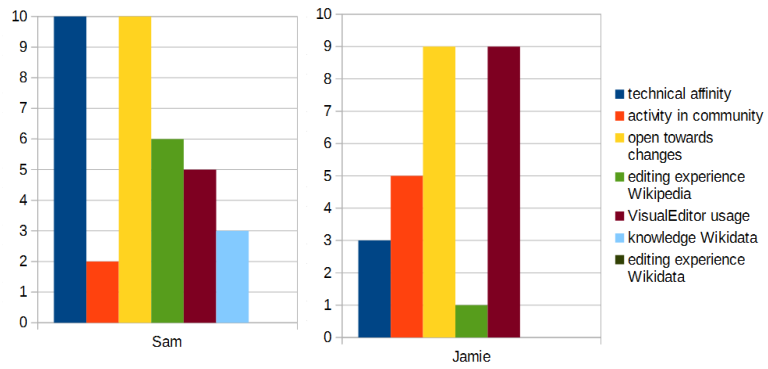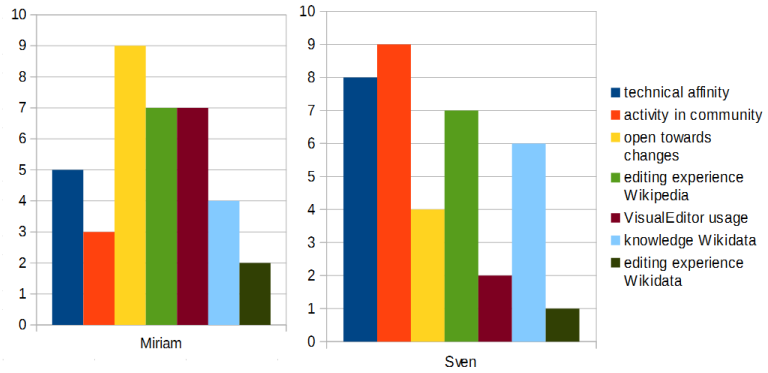
Appendix H

Usability Test Evaluation

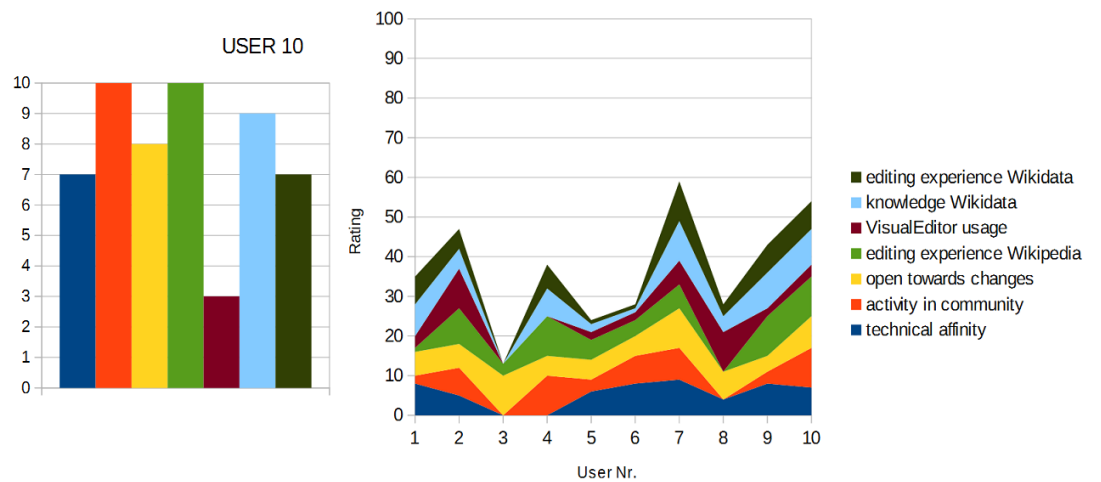**Graphs sorted by questions:**

Technical Affinity

Activity in Community

Open Towards Change

VisualEditor Usage

Wikidata Knowledge

Graphs sorted by Personas:

Graphs sorted by Usability Test Participants:

USER 2

USER 3

- technical affinity
- activity in community
- open towards changes
- editing experience Wikipedia
- VisualEditor usage
- knowledge Wikidata
- editing experience Wikidata

USER 4

USER 5

- technical affinity
- activity in community
- open towards changes
- editing experience Wikipedia
- VisualEditor usage
- knowledge Wikidata
- editing experience Wikidata

USER 6

USER 7

- technical affinity
- activity in community
- open towards changes
- editing experience Wikipedia
- VisualEditor usage
- knowledge Wikidata
- editing experience Wikidata

USER 8

USER 9

- technical affinity
- activity in community
- open towards changes
- editing experience Wikipedia
- VisualEditor usage
- knowledge Wikidata
- editing experience Wikidata

USER 10

- editing experience Wikidata
- knowledge Wikidata
- VisualEditor usage
- editing experience Wikipedia
- open towards changes
- activity in community
- technical affinity

Rating

User Nr.

Appendix I

Usability Test

1. Was siehst du?   Nr:_____
1.1 Nur gucken

1.2 Rumklicken, aber nichts editieren

Aufgaben:
1. Hole den Wert „birth date" aus Wikidata

2. Editieren den „birth date" wie gewohnt, lokal.

3. Füge eine Referenz zu „death place" hinzu.

3.1 editiere diese

3.2 lösche diese wieder

4. Füge „spouse" als field hinzu und fülle mit WD

4.1 gebe ihr einen qualifier

1. Was hat dir gefallen?

2. Was würdest du verändern?

3. Was hat dir garnicht gefallen?

4. Wie fandest du die tooltips?

5. Unter welchen umständen würdest du davon gebrauch machen?

6. andere kommentare

Checkliste:

O Wikidata button

O Apply changes to Wikidata

O die ecken

O

O

O

O

Appendix J

Self-Assessment

Nr:_____

Bevor du gehst...

Wo würdest du deine Technikaffinität einordnen?

| sehr niedrig | | sehr hoch |

Wie aktiv bist du in deiner Community?

| inaktiv | | sehr aktiv |

Wie offen stehst du Veränderungen gegenüber?

| ver- schlossen | | sehr offen |

Wo würdest du deine Editiererfahrungen auf Wikipedia einordnen?

| sehr niedrig | | sehr hoch |

Wie häufig benutzt du den VisualEditor?

| nie | | immer |

Wie gut kennst du Wikidata?

| garnicht | | sehr gut |

Wo würdest du deine Editiererfahrungen auf Wikidata eineinordnen?

| sehr niedrig | | sehr hoch |

Declaration of Authorship

I, Charlene Kritschmar, declare that this thesis titled, "Facilitating the use of Wikidata in Wikimedia projects with a user-centered design approach" and the work presented in it are my own. I confirm that:

- Where I have consulted the published work of others, this is always clearly attributed.

- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

- I have acknowledged all main sources of help.

- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:
_____

Date:
_____