Otto-von-Guericke University Magdeburg

Faculty of Computer Science

Bachelor's Thesis

Interactive Visualization of Large Concept Lattices for Data Exploration

Author:
Johannes Filter

Advisors:
Prof. Dr. Andreas Nürnberger
Otto-von-Guericke University Magdeburg

Prof. Dr. Ana García-Serrano
Universidad Nacional de Educación a Distancia

July 16, 2015

Abstract

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Inhaltsangabe

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Acknowledgements

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Contents

1. Introduction

The digital revolution is affecting every part of our life. Also the humanities scholars stand before a big change in their ways when huge analog collections are digitized. They have to apply computer science methods to organize and analyze huge amount of data. The term "Digital Humanities" evolved during the last 10 years which can be defined as an "intersection between the humanities and information technology" [38].

The Information Retrieval Department of the Universidad Nacional de Educación a Distancia (UNED) in Madrid (Spain) cooperates with human scholars to conduct research in the Digital Humanities. In this project, there are historical maps which have been digitized and annotated. To extract knowledge from the collection the research group advocates for the use of Formal Concept Analysis (FCA) for topic organization [9, 10].

They successfully implemented a FCA algorithm but lack an interactive user interface which will be developed in this thesis.

While FCA is a mathematically well-funded principle, the resulting traditional visualization of large concept lattices are a problem. Large concept lattices arose when you apply FCA to large amount of entities (Details will be explained in Chapter 2). When applying FCA to a document collection, you are likely to occur huge amount of entities. That is why alternative visualization techniques are important to get the insights of FCA, even if the lattice is large.

This thesis does not focus on the visualization of concept lattices itself. It focus on the visualization of concept lattices to explore and browse the data. This is different because it sets the data itself into the center and not the structure that was built around the data, the concept lattice.

This user interface will be web-based, to avoid dealing with setup problems you encounter when dealing with a heterogeneous operating systems environment. Because of the fast-changing environment of the web, it is important to keep with the latest technologies and techniques to not fall apart. Besides others, the software utilizes the frameworks d3.js and Bootstrap to create a pleasant user interface. The website is fast-responding because it reduces the communication between browser and web server to a minimum. In most cases, instead of reloading the page, the interface only changes DOM elements.

The remainder of this thesis is structured as follows: The background of Formal Concept Analysis will be presented in Chapter 2 and the background of User Search Interfaces in Chapter 3. In Chapter 4 I will present my (first) approach and the implementation, which will be evaluated in Chapter 5. Built on the Evaluation, I will adjust my work and present an updated version of my work in Chapter 6. I conclude in Chapter 7 and give ideas for future work.

2. Background

Before we can analyze already existing work in this area, draw our conclusions and develop our own system, we have give background information. This chapter gives an introduction into formal concept analysis, followed by an introduction into user interface design principles.

2.1 Formal Concept Analysis

Formal Concept Analysis (FCA) is a mathematically well-funded technique to analyze data. FCA builds domain ontologies from object specified by attributes. It derives from old philosophical ideas and was formalized by Rudolf Wille [20].

In the first section a small introduction is given and the second part the application for information retrieval are explained.

2.1.1 Definition

FCA is constructed from a formal context. A *formal context* is defined as as a tripple $K = (G, M, I)$ where G is a set of objects, M is a set of attributes and I is a binary relation $I \subseteq G \times M$. I specifies whether an object has an attribute or not. (G and M come from the German words 'Gegenstand' and 'Merkmal'.)

Table 2.1 illustrates an example (from David Eppstein [18]) where G comprises the integers from 1 to 10 and M comprises the attributes composite, even, odd, prime and square.

Let the operator $'$ for $A \subseteq G$ be defined as following:

$$A' = \{m \subseteq M \mid I(g, m)\ \forall g \in A\}$$

Table 2.1: Formal context, integers 1 to 10 as objects and attributes

	composite	even	odd	prime	square
1			×		×
2		×		×	
3			×	×	
4	×	×			×
5			×	×	
6	×	×			
7			×	×	
8	×	×			
9	×		×		×
10	×	×			

A' is the set of those attributes that are present in all objects from given A.

Let the operator $'$ for $B \subseteq M$ be defined as following:

$$B' = \{g \subseteq G \mid I(g, m)\ \forall m \in B\}$$

B' is the set of objects that have at least the attributes given in B.

If for $A \subseteq G$ such that $A = A''$, then A is called *closed*. The same is true for $B \subseteq M$ and $B = B''$.

For example, let a set of objects be defined as $A_1 = \{1, 4\} \subseteq G$. This results into: $A_1' = \{square\}$ and $A_1'' = \{1, 4, 9\}$. A_1 is not closed but $A_2 = 1, 4, 9 \subseteq G$ is called close because $A_2 = A_2''$.

A *formal concept* is a pair of (A, B) where $A \subseteq G$ and $B \subseteq M$ and $A = B' \wedge B = A'$. Informally, all objects in A share exactly the same attributes in B. A is a set of objects called the *extent* of a formal concept. B a set of attributes called the *intent* of a formal concept. The extent and the intent of all formal concepts are always closed.

From the example in 2.1, we can derive several formal concepts. Three randomly chosen concepts are shown in table 2.2.

Table 2.2: Three formal concepts from the formal context in table 2.1

Concept	Extent	Intent
C_1	{4,6,8,10}	{composite, even}
C_2	{2,4,6,8,10}	{even}
C_3	{9}	{composite, odd, square}

It is possible to define an order relation on the formal concepts. Let us introduce the relation $\leq$ as follows:

$$(A_i, B_i) \leq (A_j, B_j) \Longleftrightarrow A_i \subseteq A_j$$

With the help of $\leq$, we can derive relationships from the the concepts in table 2.2. We see that $C_1 \leq C_2$. This means that C_1 is more specific than C_2 and C_2 is more general than C_1. We can also see that C_3 is unrelated to C_1, and that C_3 is unrelated to C_2.

A formal context with $\leq$ is called a *concept lattice* of the context. It can be shown, that for two formal concepts C_i and C_j, there always exist a formal concept C_x such that $C_i \leq C_x \wedge C_j \leq C_x$. So there exist always exist a formal concept wich is 'above' in the hierarchy and also related to the two formal concepts.

The interested reader is advices to read "The Basic Theorem on Concept Lattices" for instance in [6] for an formal explanation.

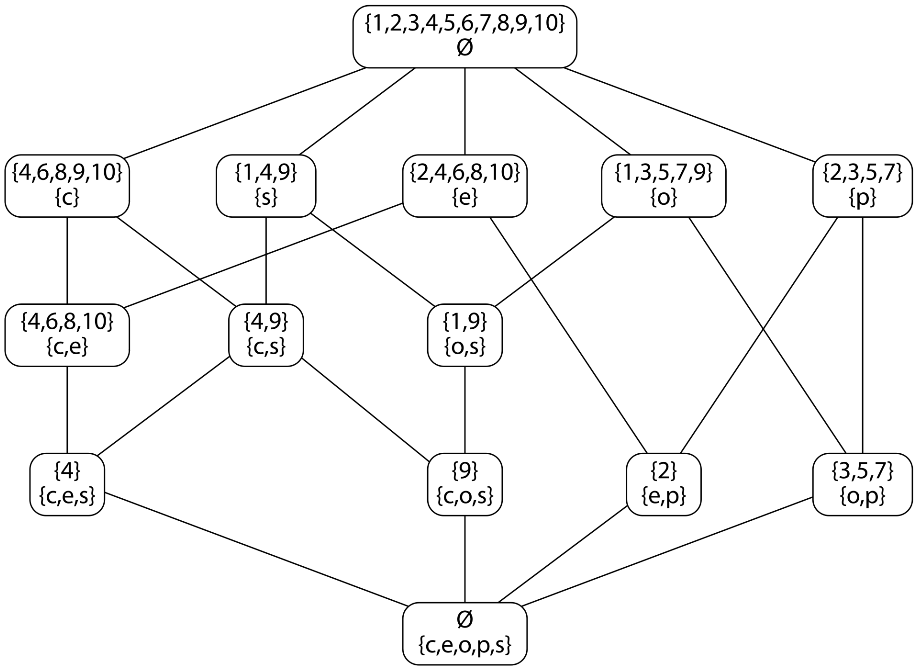
In the next section we will take a look at the static visualization of concept lattices.

2.1.2 Static Visualization

It is often said that a picture is worth a thousand words. To convey the abstract insights of a concept lattice, it can be visually represented in a *Hasse diagram*. Figure 2.1 shows the Hasse diagram of the concept lattice derived from the formal context described in table 2.1.

A Hasse diagram is a special graph where the vertices represent formal concepts and edges represent the relation $\leq$ among the formal concepts. An

Figure 2.1: Hasse diagramm, with the integers 1 to 10 as objects and attributes square (s), prime (p), composite (c), even (e), and odd (o)



egde between formal concepts C_i and C_j is drawn, when $C_i \leq C_j$ and there does not exist a formal concept C_x such as $C_i \leq C_x \leq C_j$.

. e vertices represent the formal concepts. The edges are there if they are directly related in the partial order. The most general formal concepts are in the top and the most specific ones are in the bottom. There have been added two special formal concepts. One for the most specific on the top containing all possible objects, and one most general containing no objects in the bottom.

We will describe in the next section how we can apply FCA to Information Retrieval.

2.1.3 Application for Information Retrieval

Carpineto et al.[8] describe the start of FCA in information retrieval:

> In the 80's, basic ideas were put forth - essentially that a concept can be seen as a query (the intent) with a set of retrieved

documents (the extent).

This essentially means the appliance of the Standard Boolean Retrieval Model, which, as Manning et al. [29] describe, "is a model for information retrieval in which we can pose any query which is in the form of a Boolean expression of terms [..]. The model views each document as just a set of words." . In a simple case, a system only supports conjunction of terms.

According to Poelmans et al. has been FCA "applied in many disciplines such as software engineering, knowledge discovery and information retrieval" [33] and they did two comprehensive surveys on the application of FCA [33, 34]. This thesis focusses on the visualization of a concept lattice and not on the act of creating of one. The interested reader can read literature from Carpineto et al. [6, 8] to further investigate this area.

Because the visual perspective is about for this work, let us review some design principles before discussing related work.

2.2 Interface Design

The interaction from the user with the system is what exactly matters to the user. The interaction of humans with computers own research area (human-computer interaction) and one of its pioneers is Ben Shneiderman. In the following, two principles from him will be presented: The "Eight Golden Rules of Interface Design" and the "Visual Information Seeking Mantra".

2.2.1 Eight Golden Rules of Interface Design

This rules are genreal advices for user interface designers which should apply to all interfaces. Ben Shneiderman et al. present this rules in their book [36]. The rules are explained with own remarks.

- Strive for consistency: Use similar actions in similar situations. Use identical terminology, colors, fonts etc. throughout the system.
- Cater to universal usability: Design for the needs of a diverse user group (skill level, age, gender)
- Offer informative feedback: Give system feedback for every action.
- Design dialogs to yield closure: Sequences of actions should be grouped. Give feedback on completion of a group.

- Prevent errors: Design the system that the user cannot even do errors in the first place. But if she does some, offer instructions how to recover.

- Permit easy reversal of actions: Actions should be undone. This gives the user confidence to explore the system.

- Support internal locus of control: The user should think that she is in charge of control.

- Reduce short-term memory load: Reduce the number of things the user has to keep in mind while using the system.

As you can see, the rules are open to interpretation. There exits alternative principles for instance: Donald Norman's Design Principles [32] or Jakob Nielsen's "10 Usability Heuristics for User Interface Design" [31].

This principles can be applied to all user interfaces. In the next section, design principles will be presented where the user views a large collection of items.

2.2.2 Visual Information Seeking Mantra

The visual information seeking mantra (the Mantra) was introduced by Ben Shneiderman [35] and are based on his experience with past projects. Albeit the Mantra was intended to be a "descriptive and explanatory" [3], "in effect, the Mantra has become a prescriptive principle for many information visualization designers", write Craft and Cairns [11].

The Mantra describes user interface design principles for systems when the "users are viewing collections of items, where items have multiple attributes" [35]. The starting principles are: overview first, zoom and filter, then details on demand. They will be explained below and added by three other principles.

- Overview: Gain an overview of the entire collection.

- Zoom: Zoom in on items of interest

- Filter: Filter out uninteresting items.

- Details-on-demand: Select an item or group and get details when needed.
- Relate: View relationships among items.
- History: Keep a history of actions to support undo, replay, and progressive refinement.
- Extract: Allow extraction of sub-collections and of the query parameters.

Some task needs more explanation which I will give in the following with help from related literature.

Zoom and Filter

This task are responsible for reducing the complexity of the data collection. 'Zoom' means that the user focuses on items she wants to see. 'Filter' means that she can hide items which are not interesting for her.

History

It is important to give the user the possibility to easily recover from mistakes. In addition, "it is rare that a single user action produces the desired outcome. Information exploration is inherently a process with many steps, so keeping the history of actions and allowing users to retrace their steps is important." writes Shneiderman [35].

Extract

Once interesting objects are found, the user should have the possibility to extract them from the system. Shneiderman describes printing, emailing or saving the item to the disk as 'extraction'.

2.2.3 Final Remarks

Only two, very famous design principles were presented here. The are a lot of different alternatives to choose from. There is a lot of interpretation of those guidelines involved. In addition, every system is different and has different requirements. So it is important to threat them with care, but they can help to develop an interface.

The presented ideas base mostly on the experience of one person: Ben Shneiderman. The huge number of citations show that his work is influential for a lot of people. But this is not real science. Craft and Cairns [11] are calling for empirical justification of the Mantra. This is an indication that human-computer interaction is only at the start point - there is a still a lot of research to do.

3. Related Work

With the background in FCA and interface design principles, we review and discuss related work.

3.1 Hasse Diagramm

Hasse diagrams were introduces

While some user studies proclaim that non-experts can read Hasse diagram[16], the study has been conducted on relatively small lattices. On the field of information retrieval the objects can easily outreach a few dozens. Kuznetsov et al. [26] describe this resulting visualization.

> Representing concept lattices constructed from large contexts often results in heavy, complex diagrams that can be impractical to handle and, eventually, to make sense of.

Especially enormes edge crossing can hinder the visual representation. Take a look at the appendix for the first results of the research group.

The visual representation of Hasse diagrams can be improved by fine-tuning visual components like labels, edges etc.. Or some ideas like a Fish-Eye Views XXX-QUOTE have to applied to FCA. But this actions does not scale well and wont help us with large concept lattices. To cope with large lattices, three reduction techniques exists which will be presented in the following: One where you visualize only a part of the lattice, one to transform it into a tree and one where you remove nodes from the concept lattice which means, that you modify the structure of the lattice.

3.2 Local View

Instead of showing the whole Hasse diagram to the user, only a small part of the lattice is visualized. The focus lies on one concept and its neighborhood. There exists several names and small variations of this ideas. Eklund et al. name this idea *conceptual neighborhood* [15, 17]. The user can query the system or navigate through the lattice by going up (removing terms) or going down (adding terms). Only adjacent nodes are displaced in this model. The user can incrementally browse the whole lattice. Eklund et al. applied this approach to a broad range of topic: for the 'Virtual Museum of the Pacific' [15, 17], image browsing [14, 13] and search engines [12].

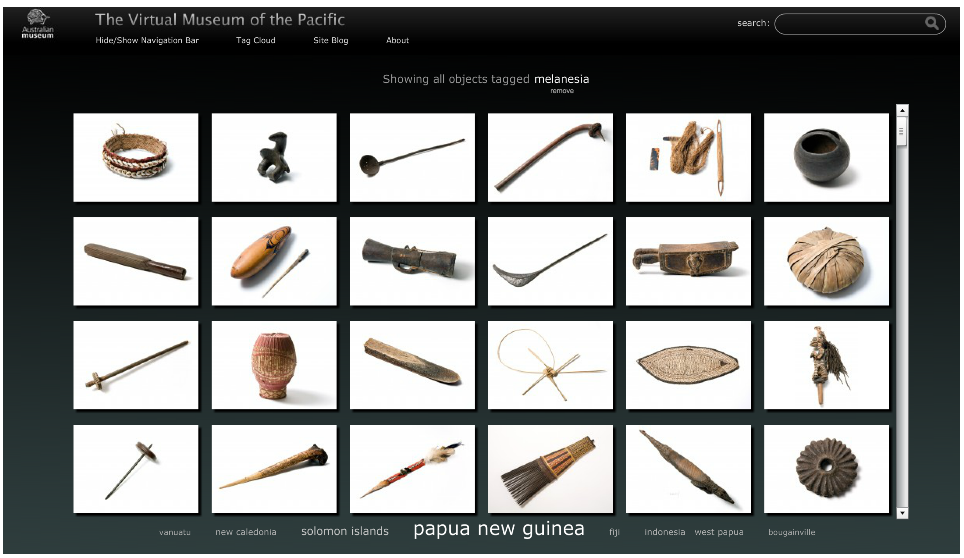


Figure 3.1: Screenshot of 'Virtual Museum of the Pacific', focus on concept 'melanesia'

Carpineto and Romano developed a search engines ULYSSES [4, 5], which visualizes the results in a similar way. But it visualizes a small sublattice - more than just the directly adjacent nodes. The size of the sublattices varies and can be fine-tuned by parameters. For instance, you can the degree of children or parents to visualize, which is the minimal distance between two nodes. You can see a screenshot of the software in Figure**??**.

In their following work, CREDO [7], Carpineto and Romano they restricted the system to only show directly neighboring nodes which a folding mechanism.

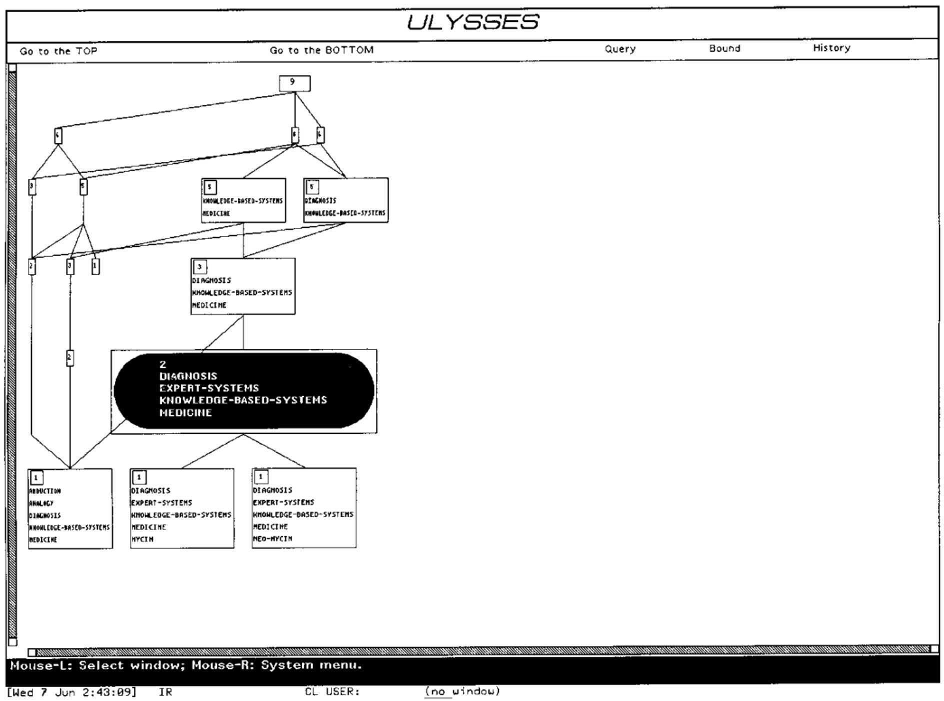


Figure 3.2: Display screen of ULYSSES, focusing on the black node [5]

3.3 Transform to Tree

While transforming the lattice into a tree sounds promising, because you could apply sophisticated tree visualizing techniques to reduce edge crossing, it comes with several drawbacks. One naiv approach is described by Carpineto and Romano [6]: If a node hast more than one parent, remove that parent an insert a copy of then node and attach it to that parent. This is problematic, because we dramatically increase the number of nodes.

There exist another approach [30]: select the 'best' parent and hide edges to all other parents. While this technically not breaks the concept, the visual representation does not correspond to the underlying model.

3.4 Pruning Nodes

A prominent approach to prune lattices called "iceberg lattice" [37]. A variation from the frequent item-set mining which specific min-support and min-confidence[1]. It creates a top of the lattice but has some drawbacks

Figure 3.3: Screenshot of CREDO, after query 'xml' and browsing after 'editor(5)' and 'windows free(2)' [7]

because "One should be careful not to overlook small but interesting groups, for example, "exotic" or "emergent" groups not yet represented by a large number of objects, or, groups that contain objects who are not members of any other group."[26] The iceberg lattice just focuses on the concepts that contain a lot of documents. That is why an other approaches exist: Stability[27]: "A concept is stable if its intent does not depend much on each particular object of the extent." [26]

It is also possible to apply traditional cluster techniques to FCA.[2]

3.5 Conclusions

We showed that there exist different possibilities to reduce the complexity of larges lattices. But in our case, we cannot change the structure of the lattice. We apply FCA to explore the data and get insights about it. When pruning the nodes, you are losing many data relationships, many formal concepts and, consequently, the "power" of FCA as exploratory technique is significantly reduced. The application of the local view technique in combination with a query interface is best for our needs. Especially the use of a query interface is familiar to all user with the rise of web search engines. The 'transform to tree' approach without a query interface seems cumbersome and all together, it is really similar to the local view approach.

Before I critically analyze the current applications of local views in chapter XX, I want to give background information in (search) interface design principles in the next chapter.

4. Fancy FCA 1.0

4.1 My Idea

We just want to find out if our approach, treating FCA as internal data structure and offering an user search interfaces + guidance (type ahead, showing neighboring concepts), is helpful for human scholars. This can be answered with a usability study.

4.2 Implementation

I will describe important parts of the implementation here.

5. Evaluation of the User Interface

The design of an interface is highly subjective. User studies can help to evaluate an interface but computer scientists are not experts in human studies and Zobel [39] proclaims: "Far too many human studies in computer science are amateurish and invalid." Nevertheless, I tried to be as scientific as possible to conduct human studies even with strict resource limitations.

We will present fundamentals of user studies first and then describe our experiments.

5.1 Fundamentals

A small introduction into this field from the computer science perspective gives Hearst in his book on User Search Interfaces in Chapter 2. [21]. A comprehensive guide into the "Methods for Evaluating Interactive Information Retrieval Systems with User" gives Kelly [24] . The interested reader is advised read those papers because we will only scratch the surface.

I will describe the idea of the experiment first and refer to literature to illustrate my choices, I will describe the process of the experiment, and after that explain and evaluate the outcome of the results.

5.1.1 Introduction

If we walk about an evaluation, it is important to make clear what what aspects are evaluated. Hearst writes [21]:

> Search interfaces are usually evaluated in terms of three main aspects of usability: effectiveness, efficiency, and satisfaction,

which are defined by ISO 9241-11, 1998 [23] as:

- Effectiveness: Accuracy and completeness with which users achieve specified goals.
- Efficiency: Resources expended in relation to the accuracy and completeness with which users achieve goals.
- Satisfaction: Freedom from discomfort, and positive attitudes towards the use of the product.

These are the criteria that ideally should be measured when evaluating a search user interface."

It is important to distinguish between the terms 'experiment' and an 'evaluation'. Kelly [24] writes: "Evaluations are conducted to assess the goodness of a system, interface or interaction technique and can take many forms [..]. Experiments have historically been the main method for interactive system evaluation, but experiments can also be conducted to understand behavior" and she continues: "Two important characteristics of experiments are that there are at least two things being compared (e.g., system type) and that some manipulation takes place." She further writes:

> "In some types of [interactive information retrieval] studies only a single system is evaluated. This is a weaker form of evaluation since it is not possible to demonstrate how much better users perform or how different their behaviors and inter actions are since there is no point of comparison. Traditional usability tests are examples of this type of evaluation. Traditional usability tests are usually conducted with a single version of a system, with the goal of identifying potential usability problems."

In this thesis, only the usability of the system is evaluated to find usability problems - we do not conduct an experiment. Experiments would be helpful to further investigate the impact of this work but exceed this thesis.

5.1.2 Categories of User Studies

Hearst [21] categorizes user studies as follows.

Informal Usability Testing

Hearst [21] describes the process shortly as "Showing designs to participants and recording their responses". It is often used in short iterative cycles to quickly evaluate a design.

Formal Studies and Controlled Experiments

Hearst [21] says that it is a "form of controlled experiments aim to advance the field's understanding of how people use interfaces, to determine which design concepts work well under what circumstances, and why." In contrast to informal studies, it is important to isolate factors and not threat the whole system as a black box. Using eye tracker in a laboratory with 2-way windows is one example.

Longitudinal Studies

"A longitudinal study tracks participant behavior while using a system over an extended period of time, as opposed to first-time usages which are what are typically assessed in formal and informal studies"

Log Analysis

In contrast to the studies above, this focus only on logs of real user interaction. Drawing conclusions from the analysis of Google Search Queries is an example.

Bucket testing (A/B Testing)

The traffic to a particular website is split and alternative view. It is evaluated how the users of the alternative website reacts to he new site. For example: Amazon changes its search filters and evaluates if the users buy more.

As you can see, this is only a small categorization and Kelly describes in her work the different approaches more in detail. Because a complete coverage of this topic would exceed this thesis, only some parts are covered here.

5.1.3 Data Collection Techniques

There exists several techniques to collect data from participants besides from interaction logs. We will describe two more here which are inexpsenvie and do not requier special, expensive euipment or are heavily time-consuming..

Questionnaires

A questionnaire comprises a set of questions and is cheap and fast way to gather information from people. Kelly et al. [25] describe two types of questions as follows:

> Questionnaires can be comprised of closed questions, open questions or a mixture of both. *Closed questions* are questions that provide a fixed set of responses with which subjects must respond. It is common practice for usability questionnaires to include closed questions in the form of statements such as, the system was easy to learn to use. Subjects are typically provided with 5–7-point Likert-type scales for responding, where one scale end-point represents strong agreement and the other represents strong disagreement. [..] *Open questions*, on the other hand, do not provide a response set and subjects are able to provide any type of response they feel is appropriate.

Questionnaires can be done with pen-and-paper, online and in an interview session. Kelly et al. [25] conducted research on different ways to elicit responses from the participants. Their results suggest that "the post-system questionnaire takes the form of an interview for closed questions, followed by pen-and-paper or electronic mode for open questions." [25]

Hornbæk and Law did a well respected meta-analysis of usability studies and as one of their conclusions they "recommend that standard questionnaires be used when possible, given their higher reliability, and that the more complex effectiveness measures be used when feasible (as they are more likely to give information that cannot be obtained by measures in the other categories)." [22]

Thinking Aloud

Kelly [24] writes by referring to Ericsson and Simon [19]: "The think-aloud method asks subjects to articulate their thinking and decision-making as they engage in [interactive information retrieval]". The comments from the participants have to be collected. Either by recording the session or by taking notes. It is hoped that the conductors can learn from the thinking process of the participant. There exist variations like that the participant should not always report because it can be exausting, challening and akwards to report all the time. Called "Spontaneous and Prompted Self-Report" the participant is encourges to report a some points or when he wants to do it.

5.1.4 Tasks

The particpants can specific instructions what to do in the experiment. They can be very concrept or vage formulated. There exists studies which show that there is a correlatin betwwen number of different task and found design errors.

5.1.5 Participants

It is important to reduce a structural bias of an experiment. Zobel [39] mentions that "the sample of human subjects should be representative (a class of computer science students may not be typical of users of mobile devices)". We tried to vary the users or at least focus on human scholars because that is the user group that is important for our stuff.

5.2 1. Test

We stick to their advice and use the USE questionnaire [28] which was (partly) used in the investigation from Kelly et all. [25].

5.2.1 Hypothesis

The interface is useful for the human scholars and they like the interface. It offers rich possibilities for the users to navigate along the different documents. Because of the similarity to popular search engines, they know how to use it. But there some stuff that is not implemented yet that they would like to see.

5.2.2 Evaluation Design

Because of our restricted resources, an informal usability study is the most attractive choice for us. The logs of the system are recored and can evaluated in future, but because of the long-term duration, they cannot be evaluated in this thesis.

5.3 Results

5.4 Conluciosn

6. Fancy FCA 2.0

7. Second User Evaluations

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

8. Conclusions

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Bibliography

[1] Rakesh Agrawal, Tomasz Imieliński, and Arun Swami. Mining association rules between sets of items in large databases. *ACM SIGMOD Record*, 22(2):207–216, 1993.

[2] Ch Aswani Kumar and S. Srinivas. Concept lattice reduction using fuzzy K-Means clustering. *Expert Systems with Applications*, 37(3):2696–2704, 2010.

[3] Stuart K Card, Jock D Mackinlay, and Ben Shneiderman. Readings in Information Visualization: Using Vision to Think. In *Information Display*, volume 1st, page 686. 1999.

[4] Claudio Carpineto and Giovanni Romano. ULYSSES: a lattice-based multiple interaction strategy retrieval interface. *Human-Computer Interaction*, pages 91–104, 1995.

[5] Claudio Carpineto and Giovanni Romano. Information retrieval through hybrid navigation of lattice representations. *International Journal of Human-Computer Studies*, 45(5):553–578, 1996.

[6] Claudio Carpineto and Giovanni Romano. *Concept data analysis: Theory and applications*. John Wiley & Sons, 2004.

[7] Claudio Carpineto and Giovanni Romano. Exploiting the potential of concept lattices for information retrieval with CREDO. *Journal of Universal Computer Science*, 10(8):985 – 1013, 2004.

[8] Claudio Carpineto and Giovanni Romano. Using Concept Lattices for Text Retrieval and Mining. *Formal Concept Analysis*, pages 161–179, 2005.

[9] Ángel Castellanos, Ana García-serrano, Etsi Informática Uned, Juan Cigarrán, and Etsi Informática Uned. Concept-based Organization for

semi-automatic Knowledge Inference in Digital Humanities : Modelling and Visualization.

[10] J Cigarrán and A Castellanos. A step forward in Topic Detection Algorithms : An Approach based on Formal Concept Analysis . 16.

[11] Brock Craft and Paul Cairns. Beyond guidelines: What can we learn from the visual information seeking mantra? *Proceedings of the International Conference on Information Visualisation*, 2005:110–118, 2005.

[12] Frithjof Dau, Jon Ducrou, and Peter Eklund. Concept similarity and related categories in SearchSleuth. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5113 LNAI(October 2007):255–268, 2008.

[13] Jon Ducrou and Peter Eklund. an Intelligent User Interface for Browsing and Searching Mpeg-7 Images Using Concept Lattices. *International Journal of Foundations of Computer Science*, 19(02):359–381, 2008.

[14] Jon Ducrou, Björn Vormbrock, and Peter Eklund. FCA-based browsing and searching of a collection of images. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4068 LNAI:203–214, 2006.

[15] P. Eklund, P. Goodall, T. Wray, B. Bunt, a. Lawson, L. Christidis, V. Daniel, and M. Van Olffen. Designing the digital ecosystem of the Virtual Museum of the Pacific. *2009 3rd IEEE International Conference on Digital Ecosystems and Technologies, DEST '09*, (June):377–383, 2009.

[16] Peter Eklund, Jon Ducrou, and Peter Brawn. Concept lattices for information visualization: Can novices read line-diagrams? *Concept Lattices. Lecture Notes in Computer Science*, 2961:57–73, 2004.

[17] Peter Eklund, Tim Wray, Peter Goodall, and Amanda Lawson. Design, information organisation and the evaluation of the Virtual Museum of the Pacific digital ecosystem. *Journal of Ambient Intelligence and Humanized Computing*, 3(4):265–280, 2012.

[18] David Eppstein. Hass diagramm, with the integers 1 to 10 as objects and attributes square, prime, composite, even, and odd. `https://commons.wikimedia.org/wiki/File:Concept_lattice.svg`, 2006. Accessed on 16. June 2015.

[19] K A Ericsson and H A Simon. *Protocol analysis: Verbal reports as data*, volume 23. 1993.

[20] Bernhard Ganter and Rudolf Wille. *Formal concept analysis: mathematical foundations*. Springer Science & Business Media, 2012.

[21] Marti a Hearst. Search User Interfaces. *Search User Interfaces*, 54(Ch 1):404, 2009.

[22] Hornb, Aelig, Kasper K, and Effie Lai-Chong Law. Meta-analysis of correlations among usability measures. *Proceedings of ACM CHI 2007 Conference on Human Factors in Computing Systems*, 1:617–626, 2007.

[23] ISO. Ergonomic requirements for office work with visual display terminals (vdts) - part 11 : Guidance on usability. ISO 9241-11, International Organization for Standardization, Geneva, Switzerland, 1998.

[24] Diane Kelly. Methods for Evaluating Interactive Information Retrieval Systems with Users. *Foundations and Trends® in Information Retrieval*, 3(1—2):1–224, 2007.

[25] Diane Kelly, David J. Harper, and Brian Landau. Questionnaire mode effects in interactive information retrieval experiments. *Information Processing and Management*, 44(1):122–141, 2008.

[26] Sergei Kuznetsov, Sergei Obiedkov, and Camille Roth. Reducing the Representation Complexity of Lattice-Based Taxonomies. pages 241–254, 2007.

[27] Sergei Kuznetsov, Sergei Obiedkov, and Camille Roth. Reducing the Representation Complexity of Lattice-Based Taxonomies. pages 241–254, 2007.

[28] Arnold M Lund. Measuring usability with the use questionnaire. *Usability interface*, 8(2):3–6, 2001.

[29] Christopher D. Manning and Prabhakar Raghavan. An Introduction to Information Retrieval, 2009.

[30] Cassio Melo, Bénédicte Le-Grand, Marie Aude Aufaure, and Anastasia Bezerianos. Extracting and visualising tree-like structures from concept lattices. *Proceedings of the International Conference on Information Visualisation*, 6:261–266, 2011.

[31] Jakob Nielsen. 10 usability heuristics for user interface design. *Fremont: Nielsen Norman Group.[Consult. 20 maio 2014]. Disponível na Internet*, 1995.

[32] Donald A Norman. *The design of everyday things: Revised and expanded edition.* Basic books, 2013.

[33] J Poelmans and So Kuznetsov. Formal concept analysis in knowledge processing: a survey on models and techniques. *Expert systems with ...*, (2003):1–40, 2013.

[34] J Poelmans and So Kuznetsov. Formal concept analysis in knowledge processing: a survey on models and techniques. *Expert systems with ...*, (2003):1–40, 2013.

[35] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. *Visual Languages, 1996. Proceedings., IEEE ...*, pages 336–343, 1996.

[36] Ben Shneiderman, Catherine Plaisan, Maxine Cohen, and Steven Jacobs. *Designing the User Interface: Strategies for Effective Human-Computer Interaction.* 5th edition, 2010.

[37] Gerd Stumme, Rafik Taouil, Yves Bastide, Nicolas Pasquier, and Lotfi Lakhal. Computing iceberg concept lattices with TITANIC. *Data and Knowledge Engineering*, 42(2):189–222, 2002.

[38] Patrik Svensson. The Landscape of Digital Humanities. *DHQ: Digital Humanities Quarterly*, 4(1):1–34, 2010.

[39] Justin Zobel. *Writing for Computer Science.* 2004.