Graz University of Technology
Knowledge Management Institute

On the Navigability of Social Tagging Systems

Dipl.-Ing. Christoph Trattner, BSc

DISSERTATION
in fulfillment of the requirements
for the academic degree
Doctor of Technical Sciences (Dr.techn.)
in Computer Science
at the
Graz University of Technology

Advisor:
Assoc. Prof. Dr. Dipl.-Ing. Denis Helic

Graz University of Technology, Austria

2nd Reader:
Prof. Dr. Dipl.-Ing. Michael Granitzer

University of Passau, Germany

Graz, 2012

To my family

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Place	Date	Signature

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

Ort	Datum	Unterschrift

Acknowledgments

First and foremost I would like to thank my advisors, Denis Helic, who helped and guided me throughout my Bachelor's studies, my Master's studies and this research, for the past few years. With Denis I have learned a lot of new things and together we worked on a number of significant research projects that were presented at top-tier international conferences.

Furthermore, I am also very grateful to Peter Brusilovsky, Yiling Li and Denis Parra from the School of Information Science, with whom I had the pleasure to work on my dissertation topic during my research stay in Pittsburgh, USA. The visit was helpful in many aspects, such as getting to know other researchers from the same field and also for the joint writing of interesting papers together that were presented at prestigious international conferences and published in top-tier international journals.

I also want to thank all my friends and colleagues at the Graz University of Technology who have been with me all this time. In particular, I would like to thank Hermann Maurer, Markus Strohmaier, Keith Andrews, Christian Körner, Markus Muhr, Philip Singer, Claudia Wagner, Florian Klien, Simon Walk, Silvia Mitter, Karl Voit, Dana Kaiser and Gerhard Wurzinger for supporting me with many fruitful discussions and feedback during my research the past few years.

Last but not least, I would like to thank Frank Kappe and Gerhard Pail who supported me with funding during the period of time when no research grants were available.

Finally, I would sincerely like to thank my parents, Christine and Leonhard Trattner, my sister Alexandra and my brother Stephan. I love you!

Institutional Acknowledgements

This PhD work was greatly supported by a grant from the BMVIT - the Federal Ministry for Transport, Innovation and Technology (grant no. 829590), project NAVTAG - Enhancing the Navigability of Tagging Systems. Parts of this thesis was also supported by funding from the Marshalplan Funding Agency.

Credits and Copyright

The thesis was written using Arkaitz Zubiaga's skeleton which can be downloaded from http://www.zubiaga.org/thesis/. Many thanks Arkaitz for letting me use this awesome template for the writing of my dissertation. Furthermore, I would like to thank Harald Krottmaier whose PhD thesis helped me to some extent to structure this work.

Abstract

Social tags provide an easy and intuitive way to annotate, organize and retrieve resources from the Web. Promoted by several pioneering systems such as Delicious, Flickr, and CiteULike, social tagging has emerged as one of the most popular technologies of the modern Web.

The value of tags was specifically advocated for information systems where the presence of tags made resources searchable and discoverable. While tags helped to discover content with a standard keyword-search, the most innovative feature of social tags was the ability to support browsing-based access to information through so-called "tag clouds". Effectively, tag clouds, are a new "social" way to find and visualize information providing both: one-click access to information and a snapshot of the "aboutness" of a tagged collection. Not unexpectedly, a large volume of research was devoted to algorithms for better tag cloud visualization. Surprisingly, only little research has questioned the usefulness of tags for efficient information access. To the best of our knowledge none of the previous works has studied the extent to which tags and corresponding tag-constructs are useful for *efficiently* searching for or *navigating* to the resources of an tag-based information system.

To that end, this dissertation aims at studying the utility of tags and corresponding tag-constructs for efficient search and navigation in tagging systems. We start in this field with a review of related work in this area and present thereafter two studies that aim to assess the usefulness of tags and tag clouds for the task of efficient search and navigation in tag-based information systems. After that we explore the navigational differences of tags compared to other tag-alike meta-data structures such as keywords and query terms. Finally, we introduce a number of novel approaches that focus on the improvement of the navigability of social tagging systems.

To the best of our knowledge this is the first work that extensively studies the utility of tags and corresponding constructs for efficient search and

navigation and that presents a number of novel approaches that enhance the navigability of social tagging systems.

Kurzfassung

Soziale Tags erlauben es, auf einfache und intuitive Art und Weise Ressourcen im Web zu annotieren, zu organisieren und wieder zu finden, und wurden durch mehrere Pioniersysteme wie zum Beispiel Delicious, Flickr und CiteULike zu einer der populärsten Technologien im modernen Web der vergangenen Jahre.

Der Einsatz von Tags wurde im Speziellen für Informationssysteme befürwortet, in denen die Vorhandensein von Tags das Auffindbarmachen von Ressourcen erleichterte. Einerseits halfen Tags Inhalte mit einer Standard-Schlüsselwortsuche zu finden, andererseits ermöglichten diese einen Browsing-basierten Zugang zu Informationen über so genannte "Tag Clouds" zu schaffen. Folglich werden Tag Clouds heutzutage als eine Art neuer "sozialer" Zugang angesehen, um Informationen zu visualisieren. Es ist daher nicht überraschend, dass in den vergangen Jahren viel Zeit und Energie in die Erforschung besserer Tag Cloud Visualisierungsalgorithmen gesteckt wurde, um den User dabei zu unterstüzen, Information einfacher wieder zu finden. Trotz zahlreicher Veröffentlichungen in diesem Bereich existiert nur sehr wenig Forschungsarbeit, welche belegt, dass Tags in der Form von Tag Clouds auch einen effizienten Informationszugriff erlauben.

Das Ziel der vorliegenden Dissertation ist es deshalb, den Wert von Tags und entsprechender Tag Konstrukte für den Task der effizienten Suche und Navigation in Informationssystemen zu quantifizieren. Zu diesem Zweck wird zum Einstieg in diese Dissertation eine ausführlichen kritischen Begutachtung verwandter Arbeiten durchgeführt; darauf folgend werden zwei Studien präsentiert, welche zum Ziel haben, den Nutzen von Tags und Tag Clouds für den Task der effizienten Suche und Navigation in Tag-basierten Informationssystemen zu evaluieren. Darauffolgend werden zwei weitere Studien vorgestellt mit dem Zweck, Unterschiede von Tags und verwandten Metadaten-Strukturen, wie Schlüsselwörtern oder Suchtermen, in naviga-

tionaler Hinsicht aufzuzeigen. Schließlich werden eine Reihe von neuartigen Algorithmen zur Konstruktion von Tag Clouds und Hierarchien vorgestellt und evaluiert, mit dem Hintergrund die Navigation in Sozialen Tagging Systemen zu verbessern.

Diese Dissertation befasst sich erstmals mit der Frage, inwieweit sich Tags und entsprechende Tag Konstrukte für eine effiziente Suche und Navigation in Tagging Systemen eignen, und präsentiert bzw. evaluiert neue Ansätze, welche es ermöglichen, die Navigation in Sozialen Tagging Systeme zu verbessern.

Contents

Part I

Overview and Scope of this Dissertation

1 Introduction

1.1 Motivation

Recently, with the emergence of modern Web 2.0 applications such as Delicious (http://delicious.com/) or Flickr (http://www.flickr.com/) social tagging systems gained tremendously in popularity [96]. In these systems, users are allowed to add simple keywords (=tags) without predefined vocabulary to describe or categorize the resources. A subset of tagging systems comprising the tagging bookmarking services like Delicious, CiteULike (http://www.citeulike.org/) or BibSonomy (http://www.bibsonomy.org/) have received community focus due to ease of use and information discovery mechanisms. In these systems users assign tags to the addresses (URLs) of resources, e.g. Web pages [50]. The weighted set of tags assigned to a resource by all users within a system and visualized as a navigation support is called the tag cloud. Effectively, tag clouds, are a "social" way to find and visualize information providing both: one-click access to information and a snapshot of the "aboutness" of a tagged collection. Not unexpectedly, a large volume of research was devoted to algorithms for better tag cloud visualization. Surprisingly, only little research has questioned the usefulness of tags for efficient information access. Hence, the question to what extent these browsing constructs are useful in efficient information access remain largely elusive. Even though studies have investigated the utility of tags to some extent from an information-theoretic or search-interface perspective [32, 143, 122], to the best of our knowledge there are no studies investigating the extent to which tags and corresponding tag-constructs are useful for searching for or efficiently navigating to the resources of tag-based information systems. To that end, the purpose of this dissertation is to extensively study the utility of tags (and related meta-data), tag clouds and tag hierarchies for the task of search and efficient navigation in online tag-based information systems.

To start in this work, we provide in this first chapter an introduction to the concept of tagging, tagging systems, navigation in such systems and finally outline the problem statement, the research questions and the structure of this work.

1.2 Social Tagging

Tagging describes the process to apply short strings, terms or words to digital resources, such as images, videos, books, etc. in a system [79]. These short character sequences are typically referred to as *tags* [151, 96]. Tagging is typically performed to describe or categorize content for later information retrieval [47, 81, 105, 127]. Hence, tagging is very related to the process of keyword application. While traditional classification systems rely on predefined or controlled vocabulary, tagging allows the users to assign tags in a free and unbound manner [79].

The process for tag application can be either collaborative or non collaborative, which is depended on the system that supports tagging. In a system that supports collaborative tagging many users are enabled to apply tags to any resource of the system [47], while in a non-collaborative environment only a single user can annotate certain resources.

We speak about *social* tagging if more than one user is accountable for the tag application process in the system [47]. A system that enables social tagging, is usually referred to as a social tagging system.

1.3 Social Tagging Systems

In the past, social tagging systems were typically associated with online bookmarking systems such as Delicious or CiteULike where people shared their bookmarked resources with others over the Web. With the emergence of social platforms such as Flickr or LastFM which integrated tagging functionality primarily to enrich online media content with lightweight meta-data for better information retrieval this paradigm changed. Nowadays, the term social tagging systems is usually referred to any system that provides social tagging functionality either as a main feature or as a supplementary function. In the following, we give a short overview of the most prominent tagging systems currently existing and shortly outline their domains (cf. [79]):

Online Bookmarking: The most popular tagging system is the online bookmarking site Delicious (`http://www.delicious.com`). Originally founded in 2003 as a social bookmarking service, Delicious was acquired from Yahoo! in early 2011 by AVOS, a company helmed by YouTube founders Chad Hurley and Steve Chen. The site was rebuilt from the ground up and re-launched in fall 2011 with a new focus on curation and discovery [35]. By the end of 2008, the service claimed more than 5.3 million users and 180 million unique bookmarked URLs [150]. In Delicious a tag is referred to as the label that describes the bookmark that is shared over the platform. The bookmark itself is stored in the users own library but is accessible through tags for all user's of the delicious platform.

Online Photo Sharing: Another very popular tagging system is the online photo sharing system Flickr (http://www.flickr.com/) where users are able to upload images and organize or share their photos with tags. Within Flickr tags are used to describe the contents of pictures, express feelings or opinions or to catalog photos into events [79]. Other popular online systems in this context are Google Picasa (http://picasaweb.google.com/) or 500px (http://www.500.px.com/).

Online Bibliographic Management: Other popular tagging systems, especially well-known in the scientific domain, are the online bibliographic management systems CiteULike (http://www.citeulike.org/), Bibsonomy (http://www.bibsonomy.org/) or Mendeley (http://www.mendeley.com/). They support the users in writing scientific articles and manage their references. Tags in this context are typically used for categorization and later information retrieval.

Online Music: Well known advocates of tagging systems in the music domain are the online platform Youtube (http://www.youtube.com/) and LastFM (http://www.lastfm.com/). Similar to the image domain, tags in this context are usually used to enrich the content with additional meta-data information and for better information retrieval.

Online Libraries and Stores: Another popular tagging system is the online platform LibraryThing (http://www.librarything.com/) supporting the user to organize her online book library through tags. In the domain of online stores Amazon (http://www.amazon.com/) is the most popular example of a tagging system providing its users with the functionality to annotate products with tags. Again, the main purpose of tags in this context are to enhance the information retrieval properties of these systems.

Online Social Networks: Last but not least we describe popular examples of tagging systems in the online social network domain. Online social networks typically provide tagging of their contents through the so-called hash-tags. Hash-tags (e.g. #tugraz) are usually used inline a message or a description to stress that the content belongs to a specific stream, topic or event [79]. Popular examples of such systems are for instance Twitter (http://www.twitter.com/), Google+ (https://plus.google.com/) or Instagram (http://instagr.am/).

1.4 Navigation in Tagging Systems

The main navigational structure in tagging systems are tags. Exceptions in this context are platforms such as Amazon or LastFM which integrate tagging functionality as a supplementary feature. In this platforms tags usually serve as additional navigational element. No matter how prominent tagging features are used, to navigate from resource A to resource B within a tagging system, the user clicks on a tag t_A applied to A retrieving a list of results where the reference r_B of B is shown. By clicking on r_B the user is then referred to B (see Figure 1.1). Depending on the contents of the tagging system a resource can be either a bookmark, an image, a movie file, a text document, etc.

For a better user interface experience tags are typically displayed as tag clouds. A *tag cloud* is a browsing interface that typically shows the top most N tags of a resource (= resource-specific tag cloud), a set of resources (= related tags tag cloud) or the resources of the whole tagging system (= global tag cloud). The tags in the tag clouds are usually sorted by alphabet and boosted in their font size according to the number of times the tag was applied to the resources. Today a large variety of tag cloud calculation algorithms exist. Some of them display tags in different colors, some of them cluster tags into categories or according to their semantic meaning, while others manipulate the font size, the intensity of the tags or simply display the tags as a simple list in alphabetic order (cf. [14, 114, 54, 68, 121, 153, 51, 76]). The decision for an optimal layout is thereby driven by the expected user goals [93]. No matter how tag clouds are calculated, the process of navigation with these browsing constructs is the same as navigating with tags. By clicking on a tag in the tag cloud the user is prompted with a result list containing resources related to this tag.

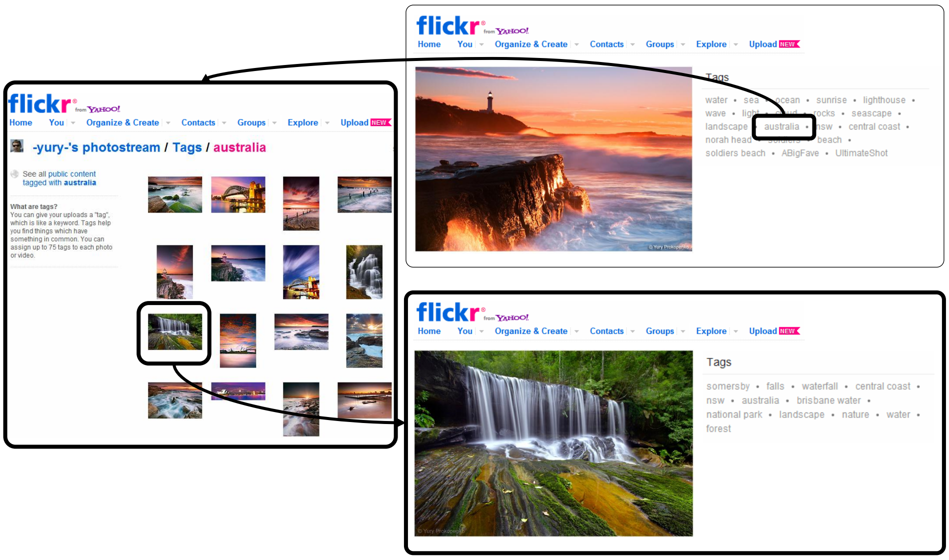
In Figure 1.1 we present an example of tag-based navigation in Flickr and Delicious. While Flickr displays the tags in resource-specific tag clouds, Delicious presents tags in a list view.

1.5 Problem Statement and Research Questions

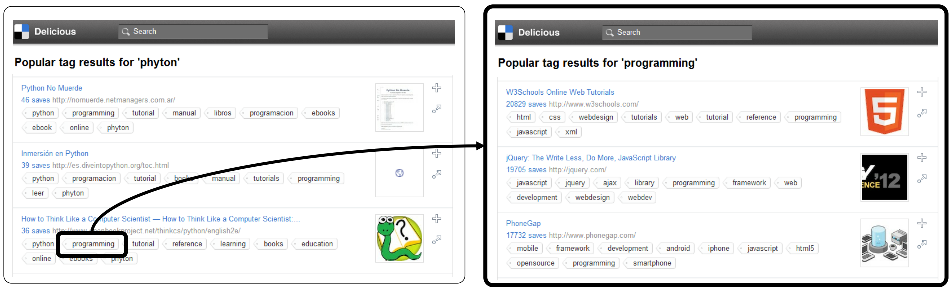
This work aims to investigate the utility of tags for the task of search and navigation in social tag-based information systems formally introduced as social tagging systems. While related work in this area has mainly focused on the visual aspects of tags [14, 114, 54, 68, 121, 7, 153, 51, 76] or has investigated tags to some extent from an information-theoretic [32, 143] or search-interface perspective [122], to the best of our knowledge none of the previous works has studied the extent to which tags are useful in efficiently searching and navigating the resources of a tagging system.

Problem Statement

> *The problem we are facing in this dissertation is the lack of knowledge about the usefulness and the efficiency of tags and corresponding state-of-the-art tag-constructs such as tag clouds for the task of search and navigation in tagging systems.*



(a) Tag-based navigation in Flickr



(b) Tag-based navigation in Delicious

Figure 1.1: Examples of tag-based navigation.

Thus, given this problem statement, the following research questions were defined:

Research Question 1

To what extent are tags/tag clouds useful for efficient navigation in tagging systems?

The first research question we ask in this dissertation is the issue to what extent tags are useful for efficient navigation in tagging systems. Since related work has only partly answered this question from an information-theoretic perspective on one single tag dataset [32], we are interested in examining this question at a much deeper level. For that purpose we study the utility of tags from a network-theoretical perspective and overall three different large-scale tag datasets, to show whether or not tags comprise efficient navigational properties. While studying tags from a navigational perspective, we also rise the question to what extent tag clouds are useful for navigation.

Research Question 2

To what extent are tags/tag clouds useful for search?

After studying the utility of tags and tag clouds for the task of navigation in tagging systems, we are interested in answering the question to what extent tags/tag clouds are useful for the task of search in tagging systems. Since related work in this area is rare [122] and has mostly answered this question from a information-theoretic perspective we perform a extensive user study that explores the usefulness and performance of tags in search interfaces.

Research Question 3

To what extent are tags/tag clouds more useful/efficient for search/navigation than other tag-alike meta-data such as keywords or search query-terms?

Another question which we are interested in to answer in this dissertation is the question to what extent tags are more useful for navigation than tag-alike meta-data such as keywords or query terms. Since tags are very related to the notation of keywords and since related research has shown that tags are in their structure comparable to the so-called query tags harvested from search query logs [15], we are interested in studying the navigational similarities or differences of tags compared to these tag-alike meta-data structures.

Research Question 4

To what extent can we build better tag-based browsing constructs that support efficient search/navigation in tagging systems?

Since our research on tag-based browsing showed that tag clouds are limited in their functionality to support efficient search and navigation of the resources of a tagging system, we are interested in answering the question to what extent better tag-based browsing constructs can be constructed that support efficient search and navigation in tagging systems.

1.6 Organization of this Dissertation

This dissertation is based on a number of articles published/presented in several international journals/conferences. In the following, papers which are included in this dissertation are listed:

1. Helic, D., Trattner, C., Strohmaier, M. and Andrews, K. 2010. *On the Navigability of Social Tagging Systems.* In Proceedings of the Second IEEE International Conference on Social Computing (SocialCom 2010), Minneapolis, Minnesota, USA, pp. 161-168.

2. Trattner, C., Lin, Y., Parra, D., Yue, Z., Real, W. and Brusilovsky, P. 2012. *Evaluating Tag-Based Information Access in Image Collections.* In Proceedings of the 23rd ACM Conference on Hypertext and Social Media (HT 2012), ACM, New York, NY, USA, pp. 113-122.

3. Trattner, C. 2011. *Linking Related Content in Web Encyclopedias with search query tag clouds.* In the International Journal on WWW/Internet, Volume 9, Issue 2 (IJWI), pp. 33-55.

4. Helic, D., Körner, C., Granitzer, M., Strohmaier, M. and Trattner, C. 2012. *Navigational Efficiency of Broad vs. Narrow Folksonomies.* In Proceedings of the 23rd ACM Conference on Hypertext and Social Media (HT 2012), ACM, New York, NY, USA, pp. 63-72.

5. Trattner, C., Helic, D. and Strohmaier, M. 2011. *On the Construction of Efficiently Navigable Tag Clouds Using Knowledge From Structured Web Content.* In the Journal of Universal Computer Science (JUCS), Volume 17, Issue 4, pp. 565-582.

6. Trattner, C. 2011. *Improving the Navigability of Tagging Systems with Hierarchically Constructed Resource Lists: A Comparative Study.* In Proceedings of the 33rd International Conference on Information Technology Interfaces (ITI 2011), IEEE, Cavtat / Dubrovnik, Croatia, pp. 173-178.

7. Trattner, C., Körner, C. and Helic, D. 2011. *Enhancing the Navigability of Social Tagging Systems with Tag Taxonomies.* In Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies (I-Know 2011). ACM, New York, NY, USA, pp. 18:1-18:8.

8. Trattner, C. 2011. *Improving the Navigability of Tagging Systems with Hierarchically Constructed Resource Lists and Tag Trails.* In the Journal of Computing and Information Technology (CIT), Volume 19, Issue 3, pp. 155-167.

A complete list of co-authored papers is included in Appendix B.

1.6.1 Contribution to the Papers

In the following section the author describes in detail the contributions of other researchers and his own contribution to the papers accumulated in this dissertation.

- Helic, D., **Trattner, C.**, Strohmaier, M. and Andrews, K. 2010. *On the Navigability of Social Tagging Systems.* In Proceedings of the Second IEEE International Conference on Social Computing (SocialCom 2010), Minneapolis, Minnesota, USA, pp. 161-168.

The idea for writing the paper was initiated by Denis Helic, Markus Strohmaier and the author. The experiments for writing the paper were conducted by the Denis Helic and the author. The discussions about the results and the interpretations were conducted by Denis Helic, Markus Strohmaier and the author in equal parts. The paper was mainly written by Denis Helic, the author and Markus Strohmaier and to some extent by Keith Andrews.

- **Trattner, C.**, Lin, Y., Parra, D., Yue, Z. and Brusilovsky, P. 2012. *Evaluating Tag-Based Information Access in Image Collections*. In Proceedings of the 23rd ACM Conference on Hypertext and Social Media (HT 2012), ACM, New York, NY, USA, pp. 113-122.

The idea for writing the paper was initiated by the author and was mainly discussed with Yiling Lin. The user study was conducted to two thirds by the first author and to one third by Yiling Lin and Zhen Yue. The three search interfaces used for the user study were implemented by the author. The paper was written mainly by the author. Peter Brusilovsky contributed in writing the abstract, the introduction and the conclusions. Furthermore, he supported the author and Yiling Lin with many fruitful discussions during the whole research process. Yiling Lin contributed in writing the related work section. Denis Parra conducted the statistical analysis and supported the author with writing the result section of the paper.

- **Trattner, C.** 2011. *Linking Related Content in Web Encyclopedias with search query tag clouds*. In the International Journal on WWW/Internet, Volume 9, Issue 2, pp. 33-55.

The idea for the paper was initiated by the author. All experiments as well as writing the paper were conducted by the author.

- Helic, D., Körner, C., Granitzer, M., Strohmaier, M. and **Trattner, C.** 2012. *Navigational Efficiency of Broad vs. Narrow Folksonomies*. In Proceedings of the 23rd ACM Conference on Hypertext and Social Media (HT 2012), ACM, New York, NY, USA, pp. 63-72.

The idea for writing the paper was discussed among all authors in equal parts. The author contributed by writing the paper as well as by the implementation of a prototype that supported the authors with a deeper understanding in interpreting the problem statement. The experimental results presented in the paper were conducted by the first and second author.

- **Trattner, C.**, Helic, D. and Strohmaier, M. 2011. *On the Construction of Efficiently Navigable Tag Clouds Using Knowledge From Structured Web Content*. In the Journal of Universal Computer Science, Volume 17, Issue 4, pp. 565-582.

The paper was initiated and mainly written by the author. All the experiments in the paper were conducted by the author. Denis Helic and Markus Strohmaier contributed by writing the abstract, the introduction and the conclusions part of the paper.

- **Trattner, C.** 2011. *Improving the Navigability of Tagging Systems with Hierarchically Constructed Resource Lists: A Comparative Study.* In Proceedings of the 33rd International Conference on Information Technology Interfaces (ITI 2011), IEEE, Cavtat / Dubrovnik, Croatia, pp. 173-178.

The idea for writing the paper was initiated by the author. Experiments as well as the writing of the paper was done by the author. Denis Helic and Keith Andrews supported the author with discussions about the user study procedure.

- **Trattner, C.**, Körner, C. and Helic, D. 2011. *Enhancing the Navigability of Social Tagging Systems with Tag Taxonomies.* In Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies (I-Know 2011). ACM, New York, NY, USA, pp. 18:1-18:8.

The idea for writing the paper was initiated by the author. Experiments as well as most of the writing of the paper was done by the author. The second and the third author contributed by writing the abstract, the conclusions and the related work section.

- **Trattner, C.** 2011. *Improving the Navigability of Tagging Systems with Hierarchically Constructed Resource Lists and Tag Trails.* In the Journal of Computing and Information Technology, Volume 19, Issue 3, pp. 155-167.

The idea for writing the paper was initiated by the author. Experiments as well as writing was done by the author. Denis Helic supported the author with discussions about the experimental setup.

1.6.2 Detailed Structure of this Dissertation

This dissertation at hand is detailed into 5 parts and 10 chapters. Below a brief overview of the contents of each chapter and a graphical illustration of the structure of this dissertation which can be found in Figure 1.2.

Chapter 1

Introduction

We present the motivation for this dissertation, formalize the problem statement and research questions.

Chapter 2

Related Work

We provide a survey of related work.

Chapter 3

On the Navigability of Social Tagging Systems

We perform a study on tags and tag clouds and their utility for efficient navigation in tagging systems.

Chapter 4

Evaluating Tag-Based Information Access in Image Collections

We evaluate the utility of tags and tag clouds for search in tag-based information systems.

Chapter 5

Linking Related Content in Web Encyclopedias with Search Query Tag Clouds

We introduce an approach that links related content in information systems via query term clouds and evaluate the extent to which query tags are more useful for navigation than tags collected by users.

Chapter 6

Navigational Efficiency of Broad vs. Narrow Folksonomies

We study the differences and similarities of broad and narrow folksonomies for the task of navigation.

Chapter 7

On the Construction of Efficiently Navigable Tag Clouds Using Knowledge from Structured Web Content

We introduce a novel tag cloud calculation algorithm that aims to enhance the navigability of tagging systems.

Chapter 8

Enhancing the Navigability of Tagging systems with Tag Hierarchies

We discuss the potentials and limitations of tag hierarchies for the task of navigation in tagging systems. Furthermore, we introduce a novel tag hierarchy construction algorithm that supports more efficient navigation than traditionally constructed tag taxonomies.

Chapter 9

Improving the Navigability of Tagging Systems with Hierarchically Constructed Resource Lists and Tag Trails

We introduce and evaluate a generic tag cloud construction algorithm that aims to support efficient navigation in tagging systems.

Chapter 10

Conclusions and Future Directions

We summarize the main contributions of this dissertation and present the answers to our research questions. Furthermore, future directions of this work are discussed.

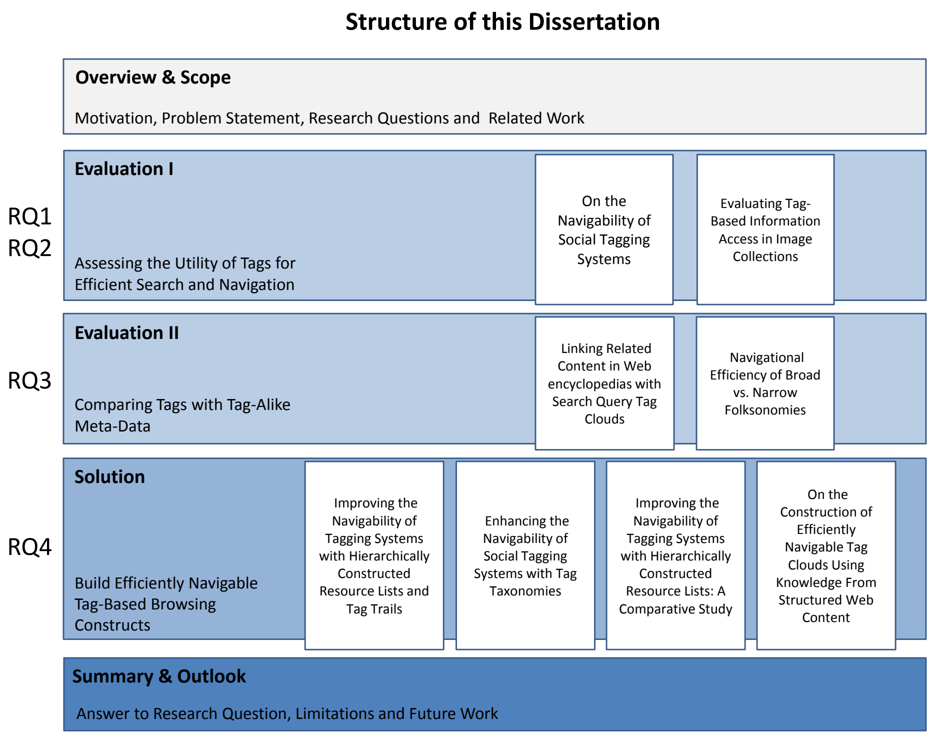


Figure 1.2: Structure of this dissertation and mapping of our contributions to the corresponding parts and research questions (cf. [79]).

2 Related Work

This chapter aims to give a broad but incomplete overview of related work. For more details, please see the corresponding related work sections of the papers included in this dissertation.

2.1 Analysis of Social Tagging Systems

The first analysis of social tagging systems was conducted by Hammond et al. In their work [50] they review nine different social tagging systems such as Delicious, CiteULike, Flickr and others and mainly analyze the features, the popularity and size of these systems.

The most cited work on the analysis of social tagging systems was conducted by Golder and Huberman in 2005. In their paper [47] the authors analyze the social bookmarking system Delicious where they discover "regularities in user activity, tag frequencies, kinds of tags used, bursts of popularity in bookmarking and a remarkable stability in the relative proportions of tags within a given URL" [47]. Furthermore, they present "a dynamical model of collaborative tagging that predicts these stable patterns in tagging systems and relates them to imitation and shared knowledge" [47].

Subsequent work by Marlow et al. [96] analyzes the tagging system Flickr and Delicious. In their work they introduce another model which gives insight into a simple taxonomy of incentives and contribution models within these systems [96].

Another impacting work in this area is the study of Halphin et al. [48]. Analyzing the complex dynamics of tagging systems they show that the distribution of the frequency of use of tags for popular sites with a long history (many tags and many users) can be described by a power law distribution [48].

The first work that analyzed tagging systems from a network-theoretic perspective is a study conducted by Shen and Wu [120]. By evaluating a crawl of the Delicious tagging system they show that tagging systems form networks that have scale-free and small world properties [120].

Another interesting work in this context is a study by Cattuto et al. In their work [31] they analyze the main network characteristics of two of these systems – Delicious and BibSonomy. To the best of our knowledge, they are the first who consider the underlying data structures of tag datasets as tripartite hypergraphs (as also considered in our dissertation) and adapt classical network measures such as characteristic path length and clustering coefficient to them [31]. Furthermore, they introduce a network of tag co-occurrence and investigate some of its statistical properties. Last but not least, they show that "simple statistical indicators unambiguously can spot non-social behavior such as spam" [31].

The last work to be mentioned in the context of social tagging system analysis is a study of Millen et al. In their paper [101] they provide results of an eight week field trial study of the enterprise social bookmarking service Dogear, including a description of user activities which were based on a log file analysis. Furthermore, their paper includes results of a survey study that focuses on the benefits and limitations of the service, including questions that reveal the usefulness of tag clouds for navigating the system [101]. Their preliminary results show a positive trend towards the usage of tags and their usefulness [101].

2.2 Studies on the Utility of Tags for Navigation

Studies on the utility of tags for navigation are rare. To the best of our knowledge there is only one study that investigates tags from this kind of perspective. In [32] Chi and Mytkowicz analyze the navigability of tags from a information-theoretic perspective and introduce the measure of entropy and conditional entropy to determine the navigability of tagging systems. By investigating the tags of the tagging system Delicious over several weeks, they find that tagging system get harder to navigate over time [32]. The reason for this is that the number of tags does not grow hand in hand with the number of resources [32]. Contrary to the work of Chi and Mytkowicz, we study in this dissertation not only one tagging system but many different tag datasets. Furthermore, we do not only focus on the navigability of tags but also focus on the evaluation of tag-based browsing constructs such as tag clouds or tag hierarchies and propose new methods and algorithms to enhance the navigability of social tagging systems. Last but not least, we study the navigability of tagging systems from a network-theoretic perspective as well as conducted user studies where possible to justify our findings.

2.3 Comparative Studies on Tags with other Kinds of Meta-Data

Another area of research related to our own work are studies that compare tags with other kinds of meta-data. Since we are interested in investigating the navigational efficiency of keywords and query terms, we list shortly related work in this area.

One of the earliest works in this context is a study conducted by Krause et al. In their research article [83] they compare tagging data from Delicious with query log data from MSN and AOL from a network-theoretic perspective. They show that "both graph structures have small world properties in that they exhibit relatively short shortest paths and high clustering coefficients" [83]. Finally, their analysis of the strength in tag-tag co-occurrence network reveals that folksonomies and logsonomies have very similar network properties [83].

In a subsequent work Benz et al. [15] extends the work of Krause et al. by additionally studying the similarities and dissimilarities of the semantic structure of folksonomies and logsonmies. They find evidence that logsonomies show a similar semantic structures as tagging data (=folksonomies) [15].

Another interesting study in this context is a work conducted by Heymann et al. In their work [63] they investigate the differences between library terms and tags. For that purpose the authors conduct a series of experiments that suggest that tagging systems tend to be at least somewhat consistent, high quality, and complete [63]. As datasets for their experiments they utilize a tag dataset from tagging systems LibraryThing and library terms obtained from the Library of Congress. Overall, Heymann et al. conclude in their work as follows: "...tags seem to do a remarkably good job of organizing data when viewed either quantitatively in comparison to "gold standard" library metadata or qualitatively as viewed by human evaluators" [63].

In this context, the probably most related work to our own one is a study conducted by Antonellis et al. In their work [9] they introduce a novel browser plug-in that generates a list of the most relevant tags and a list of the most related documents to these tags utilizing the users tags, $tf * idf$ extracted keywords or query terms. To that end they propose a novel greedy-based tag selection and link selection algorithm that maximizes the navigational gain of a Web site [9]. In a number of experiments they show that query tags are a better source for navigation than tags collected from users or keywords extracted from page-text [9].

2.4 Studies on the Utility of Tags to Enhance Web Search

Another field of research related to our dissertation are studies on the utility of tags to enhance Web Search.

In [13] Bao et al. studies social tagging in the domain of Web search. In their work they investigate the extent to which social tags from the Delicious platform can enhance Web Search [13]. They find that Web search can benefit from social tags in two aspects (1) good summarization of corresponding web pages and (2) indicator of the popularity of web pages [13].

Interesting work in this area has also been conducted by Heymann et al. in 2008. In their work [1], the authors present statistical results of a crawl of over 40 million bookmarks from Delicious and show how many bookmarks exist, how fast they grow and how useful they are to improve Web search. They find that "tags occur in over 50% of pages and in only 20% of the cases they do not occur in the page text" [1]. The conclusions they draw from these results are that tags can provide additional and meaningful data not available in other sources, though the impact for Web search may be not so high due to the lack coverage of tags over the whole Web graph [1].

In [40] Pavel et al. studies the extent to which social annotations can improve the quality of intranet search. In their work they propose two ways to obtain user annotations, using explicit and implicit feedback, and show how they can be integrated into a search engine [40]. Preliminary experiments of the authors show that social annotations improve the quality of the search results [40].

One of the most cited studies in the context of Web search and tagging is a study conducted by Bischoff et al. In their work [20] they investigate the utility of tags of three different tagging systems such as Delicious, Flickr and LastFM to search Web content. They find evidence that 50% of tags in the music domain bring new information to the resources [20]. By classifying the tags into several different categories and comparing them with data from query logs, they furthermore observe that "most of the tags can be used for search, and that in most cases tagging behavior exhibits approximately the same characteristics as searching behavior" [20].

2.5 Studies on Tag Cloud Construction and Visualization

Another question we would like to answer in this dissertation is the issue to what extent better tag-based browsing constructs can be developed to support more efficient navigation in tagging systems than current state-of-the-art approaches. In this section we list the most relevant studies related to our own work.

One of the most cited and earliest papers in this regards is a paper by Hassan-Montero and Herrero-Solana. In their article [52] they present a novel approach for tag selection in tag clouds by proposing "a clustering algorithms for visual layout, with the aim to improve the browsing experience" of the user. As evaluation metrics they introduce the notation of tag coverage and overlap. Their results show that "the presented approach reduces the semantic density of tags and improves the visual consistency of tag cloud layout" [52].

In subsequent work Kautz et al. investigate information visualization in the form of tag clusters using similarity measures such as Dice, Jaccard and Cosine similarity [77]. Based on a user study, they find that tag clusters are perceived as more useful, more trustworthy, and are more enjoyable than traditional tag cloud layouts [77].

Another study in this context is a paper by Rivadeneira et al. In their work [113] they present two studies. In the first one they compare tag layout along three dimensions: tag size, tag proximity and tag position [54]. They find "significant effects on the tag size and the location of the tags (those in the upper left are recalled better by the 13 study participants, as were those displayed with larger tags)" [54]. The second study presented in the paper includes 11 participants and a gisting task that compares four tag cloud layout algorithms with each other. Rivadeneira et al. find that "participants perform significantly better at gisting when using the simple vertical list with no font size variation" [54].

One of the largest studies in this the context of tag cloud visualization is a work by Halvey and Keane [49]. In their work they perform a user study with 62 users to investigate six tag cloud layout algorithms [49]. As evaluation method they use a selection task where users have to find a randomly chosen item within the tag clouds. The results of the study are the following [49]:

- "Alphabetization can aid users to find information more easily and quickly
- Font size is very important for how quickly and easily users find information
- The Position of tags is also very important by terms of information re-finding
- It appears that users scan lists and clouds rather than read them"

In [68] Kaser and Lemire study state-of-the-art tag cloud display algorithms and propose new models and algorithms to improve the display of tags that consist of in-line HTML, as well as algorithms that use nested tables considering tag relationships. Their results show that the proposed algorithms perform better by terms of the number of tags displayed in a tag cloud than others [68].

Another interesting work in this context is a study by Seifert et al. In their work [117] they propose a family of novel algorithms for tag cloud layout. In an an extensive user study and a technical evaluation they show the high performance of their idea. The algorithms introduced in the paper "address issues found in many common approaches, such as large whitespaces, overlapping tags and restriction to specific boundaries" [117]. "The layouts computed by these algorithms are compact and clear, have small whitespaces and may feature arbitrary convex polygons as boundaries" [117].

The last work to be mentioned in this area is a recent study conducted by Venetis et al. In their work [143] they analyze a set of tag selection algorithms that are used in current sites such as Delicious or Flickr. In order to evaluate the results of these algorithms, they introduce a synthetic user model that captures the "usefulness" of a tag cloud by terms information retrieval properties [143]. In a small user study they show that the model is a relatively good predictor for tag clouds humans prefer [143]. Contrary to our own work, they evaluate tag clouds from a information-retrieval point of view, i.e. they use tag clouds for search results summarization. In this dissertation, we are instead interested in exploring the usefulness of tags and tag clouds for the process of efficient search and navigation.

2.6 Studies on the Usefulness of Tags for Search

As outlined in the introductory part of this dissertation we are interested in studying the utility of tags and corresponding constructs for search. To that end we list the most related papers relevant to this topic in this section.

One of the earliest studies in this context is a study by Kuo et al. In their work [85], they analyze the utility of tag clouds for the summarization of search results from queries over a biomedical literature database. "The results of a user study comparing the tag-cloud summarization of query results with the standard result list provided by the system indicates that the tag cloud interface is advantageous in presenting descriptive information and in reducing user frustration" [85]. However, it is "less effective at the task of enabling users to discover relations between concepts" [85].

In subsequent work Koutrika et al. [82] introduce a framework that generates word clouds from search results trough the process of named-entity extraction. They also propose several algorithms for generating word clouds to increase the capability of the word clouds to improve search [82]. In a user study they show the high performance of their approach compared to state-of-the-art search interfaces.

One of the most prominent studies in this kind of field is the paper of Sinclair and Cardew-Hall. In their work they investigate the usefulness of tag clouds in terms of information seeking by analyzing the usage of tag clouds in a traditional search interface [122]. They find that subjects prefer tag clouds when the search task is more general, but favor issuing search queries, when more specific information is needed [122].

Interestingly, and compared to our own work none of the studies shows the utility of tags for search in two types of search behavior, look-up search and exploratory search. Furthermore, none of the studies reviles the extent to which different tag display formats affects the usefulness of tags for efficient search and information retrieval in tag-based information systems.

2.7 Studies on Extracting Hierarchical Structures from Tagging Data

Another related area of research is the work on automatic hierarchy extraction from tagging data. Since we are also interested in this dissertation in creating tag-based constructs that allow efficient navigation of the resources of a tagging system, we present in this section the most relevant papers on the extraction of hierarchical structures from tagging data, which typically represent a good way to navigate a large collection of items in an efficient manner.

One of the first studies in this context is the work of Heymann et al. [62]. In their work they show how to convert a large corpus of Delicious tags into a tag hierarchy. The groundbreaking idea of their approach is to model a tagging data as a graph and to use centrality measures to extract a tag hierarchy from this graph [62]. In subsequent work Schmitz [116] introduces an algorithm that automatically generates a taxonomy from Flickr. Contrary to the work of Heymann et al. the authors introduce an algorithm that is based on a subsumbtion-model [116]. In a number of illustrations, they show that the approach produces tag taxonomies which are semantically more sound than the hierarchies created with the algorithm of Heymann et al [116].

In [90] Li et al. introduce a novel algorithm namely Effective Large Scale Annotation Browser (ELSABer), to browse large-scale social annotation data. The novelty of the approach is the idea to generate hierarchical structures from tagging data through simple measures such as tag coverage or intersection-ratio which can be computed easily to allow efficient top-down browsing of large-scale tag datasets [90].

Plangprasopchok et al. [109] propose another hierarchy generation algorithm based on the examination of user-defined relations within a tagging system. In another work Solskinnsbakk et al. [124] constructed tag hierarchies using association rule mining of the corresponding tag set. Last but not least, in Kiu and Tsui [70] the authors introduced *TaxoFolk* - an algorithm which integrates tags and resources into a taxonomy by applying various data-mining techniques such as formal concept analysis.

Interestingly, and contrary to our work, none of these previous approaches examine the implications the resulting structures have on the navigability of the system.

2.8 Studies on Tagging Motivation and Behavior

Last but not least we briefly review recent studies on tagging behavior and motivation.

Tagging motivation was been studied by Heckner et al. In their work [55] they present results of a user study with 149 subjects to investigate the motivation for tagging in popular tagging systems including Flickr, Youtube, Delicious and Connotea. Their study reveals that tagging is performed mainly for "personal information management" and "resource sharing" [55]. Another prominent work in this area is a study conducted by Ames and Naaman examining the motivation for tagging on Flickr platform. In their work [6] they interviewed users to gain deeper insights in the motivations of people for tagging. One of the most remarkable result of their study is a taxonomy of the user's tagging motivation which has been cited more than 480 times [6].

Tagging behavior was extensively studied by Strohmaier et al. [127]. By analyzing a number of tagging datasets such as Flickr, Delicious, CiteUlike, etc. they find evidence that users can be classified into categorizers and describers [127]. In subsequent work [81] they investigate the tagging system Delicious in more detail and define a number of measures to identify these two types of user automatically.

Part II

Evaluation I: Assessing the Utility of Tags for Efficient Search and Navigation

To what extent are tags/tag clouds useful for efficient navigation? To what extent are tags/tag clouds useful for search?

3 On the Navigability of Social Tagging Systems

This chapter is based on the paper *"On the navigability of Social tagging systems"* which was presented at the Second IEEE International Conference on Social Computing in 2010.

In detail, this chapter deals with the wide-held believe that tags and tag clouds respectively facilitate efficient navigation in tagging systems. To verify this assumption we model navigation in tagging systems as a bipartite graph of tags and resources and then simulate the navigation process in such a graph. Our results show that tags/tag clouds spawn networks which are indeed efficiently navigable. However, taking user interface decisions such as "pagination" combined with reverse-chronological listing of resources into account reveals that tag clouds are significantly impaired in their potential to serve as a useful tool for navigation. Based on our findings, we identify a number of avenues for further research and the design of novel tag cloud construction algorithms.

The original contribution was published in the proceedings of the conference and can be found in [12].

3.1 Abstract

It is a widely held belief among designers of social tagging systems that tag clouds represent a useful tool for navigation. This is evident in, for example, the increasing number of tagging systems offering tag clouds for navigational purposes, which hints towards an implicit assumption that *tag clouds support efficient navigation*. In this paper, we examine and test this assumption from a network-theoretic perspective, and show that in many cases it does *not* hold. We first model navigation in tagging systems as a bipartite graph of tags and resources and then simulate the navigation

process in such a graph. We use network-theoretic properties to analyze the navigability of three tagging datasets with regard to different user interface restrictions imposed by tag clouds. Our results confirm that tag-resource networks have efficient navigation properties in theory, but they also show that popular user interface decisions (such as "pagination" combined with reverse-chronological listing of resources) significantly impair the potential of tag clouds as a useful tool for navigation. Based on our findings, we identify a number of avenues for further research and the design of novel tag cloud construction algorithms. Our work is relevant for researchers interested in navigability of emergent hypertext structures, and for engineers seeking to improve the navigability of social tagging systems.

3.2 Introduction

In social tagging systems such as Flickr and Delicious, *tag clouds* have emerged as an interesting alternative to traditional forms of navigation and hypertext browsing. The basic idea is that tag clouds provide navigational clues by aggregating tags and corresponding resources from multiple sources, and by displaying them in a visually appealing fashion. Users are presented with these tag clouds as a means for exploring and navigating the resource space in social tagging systems.

While tag clouds can potentially serve different purposes, there seems to be an implicit assumption among engineers of social tagging systems that tag clouds are specifically useful to *support navigation*. This is evident in the large-scale adoption of tag clouds for *interlinking resources* in numerous systems such as Flickr, Delicious, and BibSonomy. However, this *Navigability Assumption* has hardly been critically reflected (with some notable exceptions, for example [11]), and has largely remained untested in the past. In this paper, we will demonstrate that the prevalent approach to tag cloud-based navigation in social tagging systems is highly problematic with regard to network-theoretic measures of navigability. In a series of experiments, we will show that the Navigability Assumption only holds in very specific settings, and for the most common scenarios, we can assert that it is *wrong*.

While recent research has studied navigation in social tagging systems from user interface [24, 31, 33] and network-theoretic [28] perspectives, the unique focus of this paper is the intersection of these issues. With that focus, we want to answer questions such as: How do user interface constraints of tag clouds affect the navigability of tagging systems? And how efficient is navigation via tag clouds from a network-theoretic perspective?

Particularly, we will first 1) investigate the intrinsic navigability of tagging datasets without considering user interface effects, and then 2) take pragmatic user interface constraints into account. Next, 3) we will demonstrate that for many social tagging systems, the Navigability Assumption does not hold and we will finally 4) use our findings to illuminate a path towards improving the navigability of tag clouds.

To the best of our knowledge, this paper is among the first to study what we have called the Navigability Assumption of Tag Clouds, i. e. the widely held belief that tag clouds are useful for navigating social tagging systems. One of the main results of this paper is a more critical stance towards the usefulness of tag clouds as a navigational aid in tagging systems. We argue that in order to make use of the full potential of tag clouds, new ways of thinking about tag cloud algorithms are needed.

The paper is structured as follows: In Section 3.3 we present our network-theoretic approach to assessing navigability of tagging systems. Section 3.4 describes the analyzed datasets. Section 3.5 presents and discusses the results. Based on our findings, we call for and discuss new ideas for tag cloud algorithms in Section 3.6. Section 3.7 provides an overview of related work. Finally, Section 3.8 concludes the paper and presents directions for future work.

3.3 Network-Theoretic Model of Navigation in Tagging Systems

Typically are tagging dataset modeled as tripartite hypergraph with $V = R \cup U \cup T$, where R is the resource set, U is the user set, and T is the tag set [5, 32, 30]. An annotation of a particular resource with a particular tag produced by a particular user is a hyperedge (r, t, u), connecting three nodes from these three disjoint sets.

Such a tripartite hypergraph can be mapped onto three bipartite graphs connecting users and resources, users and tags, and tags and resources. For different purposes it is often more practical to analyse one or more of these bipartite graphs. For example, in the context of ontology learning, the bipartite graph of users and tags has been shown to be an effective projection [25].

In this paper, we focus on tag-resource bipartite graphs. These graphs naturally reflect the way users are supposed to adopt tag clouds for navigating social tagging systems. For example, in many tagging systems, tag clouds are intended to be used in the following way:

1. The system presents a tag cloud to the user.
2. The user selects a tag from the tag cloud.
3. The system presents a list of resources tagged with the selected tag.
4. The user selects a resource from the list of resources.

5. The system transfers the user to the selected resource, and the process potentially starts anew.

We will study this general interaction schema and model it with a simulated user moving along the edges of the tag-resource bipartite graph and alternately visiting tag and resource nodes.

To that end, we introduce a network-theoretic approach for assessing the navigability and the efficiency of navigability in such a bipartite graph. Ever since Milgram's *small world* experiment [26], researchers aimed to understand "navigability" and in particular "efficient" navigation of networks. Among others, two important results stem from this line of research: (1) there exist short paths between people (nodes) in a social network and (2) people are able to navigate "efficiently" through the network having only local knowledge of the network, i.e. knowing only their personal contacts.

Kleinberg [17, 16, 18] and also independently Watts [36] formalized these properties concluding that a navigable network has a short path between all – or almost all – nodes in the network [18]. Formally, such a network has a low diameter bounded *polylogarithmically*, i.e. by a polynomial in $logN$, where N is the number of nodes in the network, and there exists a *giant component*, i.e. a strongly connected component containing almost all nodes [18]. Additionally, an "efficiently" navigable network possesses certain structural properties so that it is possible to design efficient *decentralised search algorithms* (algorithms that only have local knowledge of the network) [17, 16, 18]. The delivery time (the expected number of steps to reach an arbitrary target node) of such algorithms is *polylogarithmic* or at most *sublinear* in N.

User navigation in hypertext systems is naturally modeled as a decentralised search, i.e. at each particular node in the network, users select a new node having only local knowledge of the network and following the idea that the selected node would bring them closest to their destination node. We use this model to investigate the navigability of tag clouds next.

3.4 Experimental Setup

In the following, we conduct experiments aiming to shed light on the navigability of tag-clouds in social tagging systems. We are particularly interested in studying how design decisions, such as *what tags to include in a tag cloud* or *how many tags to display*, effect the navigability of tag clouds. While, today, designers often base such decisions on intuition or heuristics, it is our goal to study the consequences of these decisions experimentally, i.e. by exploring their empirical effects on the network.

In our experiments, we used three datasets covering a range of different settings.

- **Dataset Austria-Forum**: This dataset consists of annotations from an Austrian encyclopedia called Austria-Forum[1]. The dataset contains 32,245 annotations and 12,837 unique resources. The system is at an early phase of adoption, i.e. not many users currently contribute new tags.
- **Dataset BibSonomy**: This dataset[2] contains nearly all 916,495 annotations and 235,339 unique resources from a dump of BibSonomy [14] until 2009-01-01. Annotations from known spammers have been excluded from the dataset. This dataset is obtained from a more mature tagging system.
- **Dataset CiteULike**: This dataset contains 6,328,021 annotations and 1,697,365 unique resources and is available online[3]. Again, this is a dataset acquired from a more mature tagging system.

Dataset Austria-Forum represents a tagging system at an early stage of adoption. Datasets BibSonomy and CiteULike are tagging systems which have reached a certain level of maturity (i.e. attracted a larger set of active users). While all three systems adopt tag clouds for navigational purposes, their specific approaches vary. However, because the datasets contain complete information about the tripartite graph, we can experimentally manipulate the data in a way that simulates different approaches to tag cloud construction consistently across all datasets. We will describe how we manipulate the data to simulate different user interface constraints next.

3.4.1 User Interface Issues

The first user interface restriction which we model is the size of a tag cloud, i.e. the maximal number of tags displayed in a tag cloud. While different tagging systems implement different design choices, we can simulate alternative choices across all datasets. For example, in some tagging systems the maximum number of tags in a tag cloud might be 20, while in others it might be much larger.

Another important issue of tag clouds is the algorithm used to select the tags to display in a tag cloud. While, in theory, there are many ways to compute and visualise tag clouds [8, 15, 31], in practice many tagging systems follow a simple resource-specific, *TopN* algorithm. In resource-specific approaches to tag cloud construction, only tags assigned to the corresponding resources are considered. In TopN approaches, the top n tags with the highest resource-specific frequency are chosen for display in the corresponding tag cloud. In cases where less than n tags per resource are available, the remaining slots are left empty.

[1] http://www.austria-lexikon.at
[2] http://www.kde.cs.uni-kassel.de/ws/dc09/
[3] http://www.citeulike.org/faq/data.adp

For the experiments aiming to study the Navigational Assumption, we used the TopN algorithm (because it is the most common) to reconstruct simulated networks of resource-specific tag clouds for our three datasets.

Popular tags in a mature tagging system can cover hundreds or even thousands of resources, which exceeds the pragmatic limits of a system's user interface. In this situation, tagging systems usually resort to limiting the set of resources being displayed for a given tag (for example, by sorting and "paginating" the list of corresponding resources). To model such limits, we introduce a pragmatic parameter, the length of the resource list being presented, and denote it henceforth with k.

In the majority of tagging systems, the resource lists presented after selecting a tag are usually sorted reverse-chronologically (the resources most recently tagged are listed first). For simplicity, in our experiments, we select the k resources for k-limited resource lists randomly.

3.5 Results

3.5.1 Intrinsic navigability of tagging systems

We start our study by analysing the navigability of tagging systems in a synthetic network-theoretic case, i.e. without taking any user interface restrictions into account. The first row in each of Tables 3.1a, 3.1b, and 3.1c present the obtained results. The results show the existence of a giant component connecting almost all of the nodes (98%), as well as the existence of a low effective diameter (less than 7, i.e. it is less than polynomial in $logN$, see Figure 3.1).

The only exception here is the Austria-Forum dataset. We speculate that the reason for that is due to the system being in an early adoption stage. While the effective diameter of the Austria-Forum dataset is larger than the one in the two other datasets (see Figure 3.1), it is still limited polylogarithmically, whereas the giant component contains only 77% of nodes. This result suggests that the Navigability Assumption depends on the adoption stage of the tagging system under investigation, i. e. the assumption may only hold for more mature tagging systems BibSonomy or CiteULike. We leave the issue of identifying the point in time where immature tagging systems transition to tagging systems exhibiting more useful navigational properties to future research. At this point, we simply observe that the Navigation Assumption is sensitive to the stage of adoption of a tagging system.

Result 1: The usefulness of tag clouds for navigation is sensitive to the phase of adoption of the social tagging system.

Figures 3.2a, 3.2b, and 3.2c show tag (blue), resource (green), and degree (red) distributions for the analysed datasets. The tag and resource distribu-

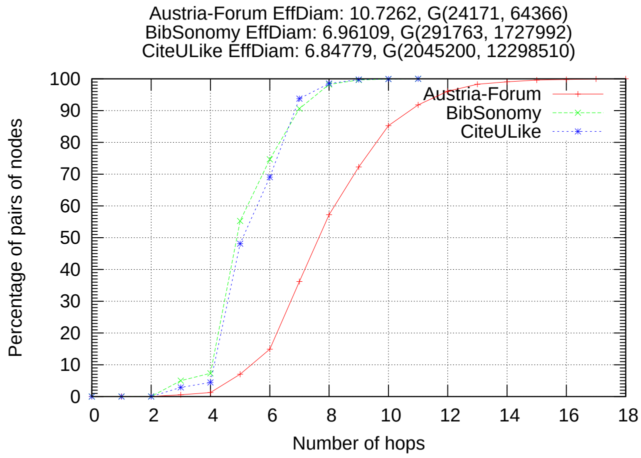
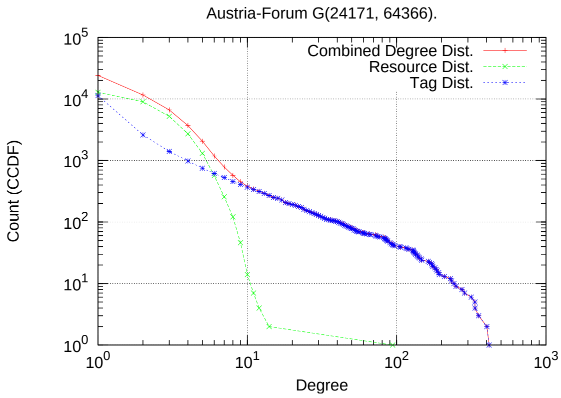


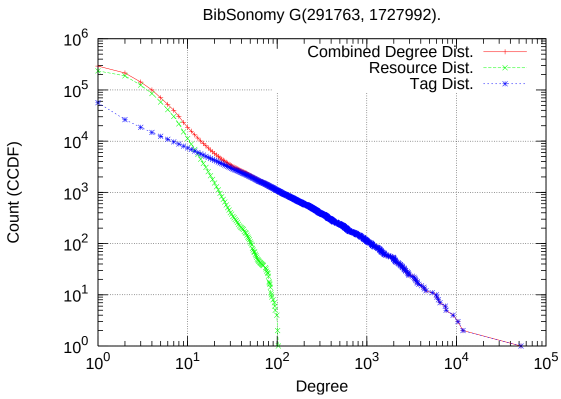
Figure 3.1: Hop plots for three different tagging datasets. We can observe the shrinking diameter phenomenon [22]: The two mature datasets (Bibsonomy and CiteULike, the two lines on the left) exhibit a small diameter, while the Austria-Forum (a tagging system in an early adoption phase, the line on the right) exhibits a larger diameter, and a larger ratio of long distances between nodes.

tions were obtained by analysing a unidirectional bipartite graph, i.e. a graph with only directed links from tags to resources. The out-degree distribution and the in-degree distribution in this graph correspond to tag distribution and to resource distribution respectively. For certain ranges of degrees, both distributions are power law distributions. There are deviations in the tail of the tag distribution – these stem from the system tags assigned to imported resources (see Figures 3.2b and 3.2c). The vertical line in the tail of Figure 3.2c comes from the existence of synonym tags in the dataset. The resource distributions exhibit an exponential cut-off in the tail (see Figure 3.2b), a deviation in the tail stemming from a test resource (see Figure 3.2a), and a power law distribution as in Figure 3.2c.

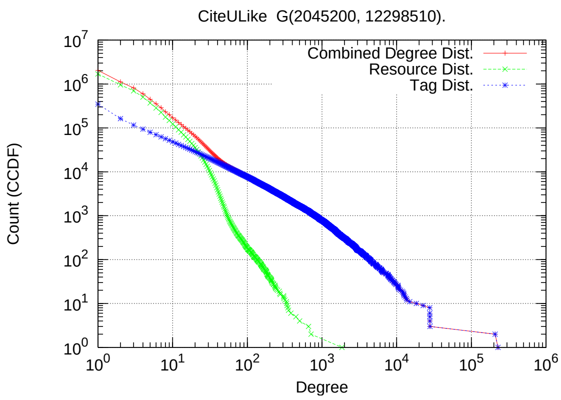
The degree distribution of the undirected bipartite graph (the red line in Figures 3.2a, 3.2b 3.2c) combines both tag and resource distributions. For lower degrees, the combined degree distribution takes the form of the resource distribution, i.e. the number of resources with low frequencies dominates the number of tags with low frequencies. For higher degrees, the combined distribution takes the form of the tag distribution, i.e. there are more tags with high frequencies than resources with high frequencies. The tag distribution is two or more orders of magnitude larger than the resource distribution, i.e. the tag distribution strongly dominates the resource distribution for higher degrees. That means that the network hubs (high-degree nodes) are the "head" tags, i.e. the top tags for *TopN* tag cloud construction algorithms.



(a) Austria-Forum



(b) BibSonomy



(c) CiteULike

Figure 3.2: Tag, resource, and degree distributions for the three datasets.

UIR	**GC**	**ED**	**UIA**	**NADT**
none	0.77	10.73	none	sub-lin.
$n = 5$	0.75	10.99	TopN	sub-lin.
$n = 10$	0.76	11.3	TopN	sub-lin.
$n = 20$	0.76	11.97	TopN	sub-lin.
$n = 30$	0.76	11.05	TopN	sub-lin.
$k = 5$	0.36	12.04	Chron.	unnav.
$k = 10$	0.47	11.16	Chron.	unnav.
$k = 20$	0.56	10.31	Chron.	unnav.
$k = 30$	0.6	10.68	Chron.	unnav.

(a) Austria-Forum

UIR	**GC**	**ED**	**UIA**	**NADT**
none	0.98	6.96	none	sub-lin.
$n = 5$	0.94	6.8	TopN	sub-lin.
$n = 10$	0.97	6.87	TopN	sub-lin.
$n = 20$	0.98	6.84	TopN	sub-lin.
$n = 30$	0.98	6.91	TopN	sub-lin.
$k = 5$	0.31	6.82	Chron.	unnav.
$k = 10$	0.4	6.62	Chron.	unnav.
$k = 20$	0.5	6.61	Chron.	unnav.
$k = 30$	0.54	6.65	Chron.	unnav.

(b) BibSonomy

UIR	**GC**	**ED**	**UIA**	**NADT**
none	0.98	6.85	none	sub-lin.
$n = 5$	0.93	6.97	TopN	sub-lin.
$n = 10$	0.95	7.07	TopN	sub-lin.
$n = 20$	0.97	7.17	TopN	sub-lin.
$n = 30$	0.97	6.98	TopN	sub-lin.
$k = 5$	0.27	6.89	Chron.	unnav.
$k = 10$	0.36	6.95	Chron.	unnav.
$k = 20$	0.44	6.91	Chron.	unnav.
$k = 30$	0.48	7.05	Chron.	unnav.

(c) CiteULike

UIR = UI Restriction, GC = Giant Component, ED = Effective Diameter, UIA = UI Algorithm, NADT = Navigation Algorithm Delivery Time Chron. = Chronological algorithm, sub-lin. = sub-linear, unnav. = unnavigable network

Table 3.1: Navigational properties of the Austria-Forum, BibSonomy, and CiteULike tagging systems.

Due to the existence of a giant component and a low diameter, tagging systems are intrinsically navigable. In [1], Adamic shows the existence of efficient decentralised navigation and search algorithms for power law networks. In principle, a user could first navigate to a hub (which is typically achieved in a few hops in a power law network) and since hubs have a large out-degree, one can reach the destination node easily. The delivery time of the algorithm is sub-linear, although the number of inspected nodes in the worst-case is $O(N)$, since sometimes the user needs to inspect all outgoing links from a hub.

Result 2: Tagging networks are navigable power-law networks. For power law networks, efficient sub-linear decentralised navigation algorithms exist.

3.5.2 Tag cloud size

Rows two to five of Tables 3.1a, 3.1b, and 3.1c show the results of applying the TopN algorithm to limit the tag cloud size on the analysed datasets. From a network-theoretic point of view, limiting the tag cloud size means limiting the out-degree of the resource nodes in the bipartite graph. The out-degree of the resource nodes is two orders of magnitude smaller then the out-degree of the tag nodes, indicating there are no resource "hubs" in the network. Therefore, limiting the tag cloud size does not influence the network to a large extent. In other words, the structure of the network is still maintained, i.e. the network remains a navigable network with navigation efficiency inherent to power law networks.

Result 3: Limiting the tag cloud size to practically feasible sizes (e.g. 5, 10, or more) does not influence navigability.

3.5.3 Pagination

Rows six to nine of Tables 3.1a, 3.1b, and 3.1c contain the results of simulating pagination with resource lists sorted reverse-chronologically. Even without experiments, it is evident that limiting the number of links going out from a tag node has destructive effects on the resulting network. In other words, limiting the out-degree of hub nodes in a power-law network destroys the connectivity of the network as a whole. Our experiments show exactly that: the giant component collapses, and the largest strongly connected component now only contains around 50% or less nodes. As such, pagination destroys network navigability, and the Navigability Assumption only holds when we assume that users would be able and willing to inspect long lists (>10.000) of resources per tag, which is not reasonable. For example, we know from search query log research that users rarely click on links beyond the first result page [38]. This yields our final result:

Result 4: Limiting the out-degree of high frequency tags (e.g. through *pagination* with resource lists sorted reverse-chronologically) leaves the network vulnerable to fragmentation. This destroys navigability of prevalent approaches to tag clouds.

3.6 Implications

The previous analysis illustrated the vulnerability of tagging networks to the *pagination effect*, where a limit is placed on the number of links going out from paginated tags, i.e. tags with frequency higher than the pagination parameter k. This vulnerability is mainly due to the simplicity of the common pagination algorithm, i.e. the resource list is simply sorted reverse-chronologically and only the k most recently tagged resources are presented to the user. The algorithm does not take into account the current user context, i.e. the resource where the user clicks on a paginated tag. Rather the same reverse-chronologically resource list is presented for a given paginated tag throughout the system.

Let us now investigate possibilities to recover the navigability of tagging networks by means of alternative tag construction algorithms. To this end, we introduce an adapted pagination algorithm. A simple generalisation of the pagination algorithm is to select k different resources out of all resources tagged with a given paginated tag, depending on the current user context, i.e. depending on the resource where the user activates a paginated tag. Let us denote the resources list of a given paginated tag t with R_t. In this case, a particular selection of resources for t becomes a function of a given resource and parameter k, i.e. $RL_t = f(r, k)$. In other words, each paginated tag is replaced by as many resource-specific tags (t_r) as there are resources in its resource list. Each resource-specific tag is then connected to resources computed by $f(r, k)$. The pseudo-code of the generalised algorithm is given in Figure 1.

We now discuss some potential functions $f(r, k)$ for selecting resources from the available resource pool and analyse their influence on network navigability.

3.6.1 Random link selection

A first obvious choice for $f(r, k)$ is to select k resources uniformly at random. This approach generates a random graph as introduced by [9] for each given paginated tag. As [4] and [3] showed, graphs generated uniformly at random are typically connected and have – with a high probability – a diameter bound by $logN$ (already for out-degrees $k \geq 3$). However, since there are no structural clues in a randomly generated network, a decentralized search

Algorithm 1 Generalized pagination algorithm

Input: $G =< V, E >$, r, t, k
for all $r \in R_t$ **do**
  add t_r to V
  add (r, t_r) to E
  $RL_t \leftarrow f(r, k)$
  for all $rr \in RL_t$ **do**
    add (t_r, rr) to E
  end for
end for
remove t from V

algorithm will need to inspect, in the worst case, all nodes of the network in order to reach a destination node from the given starting node.

Table 3.2 shows the results of a random pagination algorithm on the three test datasets. All three networks become strongly connected with a giant component even for low values of k. As expected, all three networks also possess a low diameter.

3.6.2 Hierarchical network model

In [18], Kleinberg introduced the hierarchical network model and elegantly proved that it is possible to design efficient decentralised search algorithms for such networks with a delivery time polynomial in $logN$. Put simply, Kleinberg showed that, if the nodes of a network can be organised into a hierarchy, then such a hierarchy provides a probability distribution for connecting the nodes in the network. The resulting network is efficiently navigable. A special case of the hierarchical network model is given when there is a constant number of links leaving a node, i.e. when the out-degree of a node is limited by a parameter k as it is the case with pagination. In this case, the tree leaves contain so-called *clusters* of nodes, i.e. a collection of a certain constant number of nodes.

Thus, we developed a hierarchical network generator that 1) sorts the resource list of a given paginated tag by frequency, 2) creates resource clusters of size 10 by traversing the sorted resource list sequentially, 3) creates a balanced b-ary ($b = 5$) tree where the number of leaves is equal to the number of the resource clusters, 4) traverses the tree in postorder from left to right and attaches resource clusters to the tree leaves, and 5) uses this tree structure to obtain the link probability distribution for connecting a resource-specific tag node with resources of a given paginated tag.

It is important to note that the tree creation process follows the statistical properties of the tagging dataset only, it has no inherent *semantic* rationale. As such, it serves primarily as a statistical tool to improve the efficiency

UIR	GC	ED	UIA	NADT
k=5	0.86	11.7	Random	linear
k=10	0.86	11.02	Random	linear
k=20	0.85	10	Random	linear
k=30	0.84	10.42	Random	linear

(a) Austria-Forum

UIR	GC	ED	UIA	NADT
k=5	0.99	8.75	Random	linear
k=10	0.99	6.97	Random	linear
k=20	0.99	6.75	Random	linear
k=30	0.99	6.46	Random	linear

(b) BibSonomy

UIR	GC	ED	UIA	NADT
k=5	0.99	7.98	Random	linear
k=10	0.99	7.88	Random	linear
k=20	0.99	7.13	Random	linear
k=30	0.99	6.86	Random	linear

(c) CiteULike

UIR = UI Restriction, GC = Giant Component, ED = Effective Diameter, UIA = UI Algorithm, NADT = Navigation Algorithm Delivery Time

Table 3.2: Navigational properties of the Austria-Forum, BibSonomy, and CiteULike tagging systems with a random pagination algorithm.

of navigability from a network-theoretic perspective. Table 3.3 provides an overview of the results of the structural network analysis performed with the three real-life datasets.

Another important observation is that in our model each paginated tag is a source of a network generated by a hierarchy. These networks are themselves connected through tag co-occurrence in the dataset, i.e. since tags overlap and share resources such shared resources link different generated networks. This makes it more difficult to estimate the delivery time of a decentralised search algorithm possessing only the local knowledge. If the algorithm is extended to have knowledge of all the hierarchies used in the generation of the networks, then this additional information might be useful in finding a destination node faster.

However, more theoretical work is needed to offer a proof of this intuitive assumption. In addition, it would be interesting to test these ideas empirically, for example, by implementing the algorithm and applying it to the real-life datasets. Another interesting problem is the fitting of parameters for the hierarchical network model, for example what is the optimal combi-

UIR	GC	ED	UIA	NADT
k=5	0.85	12.03	Hier.	polylog.
k=10	0.86	10.62	Hier.	polylog.
k=20	0.85	9.29	Hier.	polylog.
k=30	0.84	9.71	Hier.	polylog.

(a) Austria-Forum

UIR	GC	ED	UIA	NADT
k=5	0.99	8.82	Hier.	polylog.
k=10	0.99	7.62	Hier.	polylog.
k=20	0.99	6.94	Hier.	polylog.
k=30	0.99	6.75	Hier.	polylog.

(b) BibSonomy

UIR	GC	ED	UIA	NADT
k=5	0.99	8.76	Hier.	polylog.
k=10	0.99	7.6	Hier.	polylog.
k=20	0.99	6.36	Hier.	polylog.
k=30	0.99	5.89	Hier.	polylog.

(c) CiteULike

UIR = UI Restriction, GC = Giant Comp., ED = Eff. Diameter, UIA = UI Algorithm, NADT = Navigation Algorithm Delivery Time
Hier. = Hierarchical Algorithm, polylog. = polylogarithmic

Table 3.3: Navigational properties of the Austria-Forum, BibSonomy, and CiteULike tagging systems with a hierarchical pagination algorithm.

nation of the cluster size and the maximum number of children, with respect to the size of the resource list and the pagination parameter k.

3.7 Related Work

We start our review of related work with a brief overview of network-related research. Research on network navigability has been inspired by Milgram's *small world* experiment [26]. In this experiment, selected persons from Nebraska received a letter they were then asked to send through their social networks to a stockbroker in Boston. The striking result of the study was that, for those letters reaching the destination, the average number of hops was around 6, i.e. the population of the USA constituted a "small world". While the conclusions have been challenged [19], this experiment has attracted a great deal of interest in the research community.

Numerous researchers analysed Milgram's experiment trying to create network models and generators able to produce such "small world" networks (see for example [20]). The lattice model by Watts [37] mimics a real-life social network, where people are primarily connected to their neighbours with a few "long-range" contacts. The networks generated by this model have, like the random graph model [4, 3], a giant component and a diameter bound by $logN$.

Kleinberg analysed the second result of the Milgram's experiment, the ability of people to *find a short path* when there is such a path between two nodes [17, 16, 18]. He concluded that there are structural clues in such networks, which allow people to find a short path efficiently and argued that for an "efficiently" navigable network there exists a *decentralised search algorithm* with delivery time polynomial in $logN$.

Kleinberg also designed a number of network models such as 2D-grid models [16], hierarchical models [18], and group models [18], and showed that for certain combinations of parameters, efficient decentralised search algorithms exist. Kleinberg also showed that there is no such algorithm for the lattice model.

Particularly, hierarchical network models [18] are based on the idea that, in many settings, the nodes in a network might be classified according to a taxonomy. The taxonomy can be represented as a b-ary tree and network nodes can be attached to the leaves of the tree. For each node v, we can create a link to all other nodes w with the probability that decreases with $h(v, w)$ where h is the height of the least common ancestor of v and w in the tree. For a constant out-degree, the nodes are clustered and then the clusters are attached to the tree. The link distribution defined by $f(h) = (h + 1)^{-2}b^{-h}$ generates a navigable network with a decentralised search algorithm with delivery time of $O(log_b^4 N)$.

In related research of tagging systems, tag clouds have been characterised as a way to translate the emergent vocabulary of a folksonomy into social navigation tools [33, 7]. Social navigation itself represents a multi-dimensional concept, covering a range of different issues and ideas. A distinction between direct and indirect social navigation, for example, highlights whether navigational clues are provided by direct communication among users (e.g. via chat), or whether navigational clues are indirectly inferred from historical traces left by others [27]. Based on this distinction, our work only focuses on indirect social navigation in the sense that it studies the effectiveness of traces ("tags") left by users in tagging systems. Other types of social navigation emphasise the need to show the presence of others users, to build trust among groups of users, or to encourage certain behaviour [27].

Researchers have discussed the advantages and drawbacks of tag clouds, suggesting that tag clouds are a useful mechanism when users' search tasks are general and explorative (for example, learn about Web 2.0), while tag clouds provide little value for specific information-seeking tasks (for example, navigate to www.cnn.com) [33, 35]. While the paper at hand focuses on network-theoretic aspects, cognitive aspects of navigation have been studied previously using, for example, SNIF-ACT [10] and social information foraging theory [29]. Other work has studied the motivations of users for tagging [34, 21], and how they influence emergent semantic (as opposed to navigational) structures. The navigational utility of single tags has been investigated [6] with somewhat disappointing results. With time the tags become harder and harder to use as they lose specificity and reference too many resources. Such tags are exactly those paginated tags where new pagination algorithms are needed.

Navigation models for tagging systems have been also discussed recently. In [30] authors describe a navigation framework for tagging systems. The authors apply the framework to analyze possible attacks on tagging systems. In principle, the framework identifies a navigation channels as any combination of the basic elements of a tagging system (users, tags, and resources). Thus, the specific combination which we investigated in this paper can be summarized as the resource-tag or tag-resource navigation channel.

Recent literature also discusses algorithms for the construction of tag clouds. The ELSABer algorithm [23] represents an example of such an effort aimed towards identifying hierarchical relationships between annotations to facilitate browsing. The work by [2] is another example, introducing entropy-based algorithms for the construction of interesting tag clouds. However, these algorithms have not found wide-spread adoption in current social tagging systems. In addition, empirical studies of tagging systems have for example focused on comparing navigational characteristics of tag distributions to similar distributions produced by library terms [13].

Our work contributes to an increased theoretical understanding about the navigability of current tag cloud algorithms in social tagging systems. Our experiments identify empirical problems related to the navigability of tag clouds in three real-world tagging systems.

3.8 Conclusion

The motivation for this research was to examine and test the widely held belief that tag clouds support efficient navigation in social tagging systems. We have shown that for certain specific, but popular, tag cloud scenarios, the so-called Navigability Assumption does *not* hold. The results presented in this paper make a theoretical and an empirical argument *against* existing approaches to tag cloud construction. Our work thereby both confirms and refutes the assumption that current tag cloud incarnations are a useful tool for navigating social tagging systems. While we confirm that tag-resource networks have efficient navigational properties in theory, we show that popular user interface decisions (such as "pagination" combined with reverse-chronological listing of resources) significantly impair navigability. Our experimental results demonstrate that popular approaches to using tag clouds for navigational purposes suffer from significant problems.

Building on recent research results from network theory, in particular hierarchical network models, we have illustrated a path towards constructing more efficiently navigable tag cloud networks, which are less vulnerable to pagination influences. We conclude that in order to make full use of the potential of tag clouds for navigating social tagging systems, new and more sophisticated ways of thinking about designing tag cloud algorithms are needed.

References

[1] L. A. Adamic, R. M. Lukose, A. R. Puniyani, and B. A. Huberman. Search in power-law networks. *Physical Review E*, 64(4):046135 1–8, Sep 2001.

[2] K. Aouiche, D. Lemire, and R. Godin. Web 2.0 OLAP: From Data Cubes to Tag Clouds. *4th International Conference, WEBIST 2008*, 18, 2008.

[3] B. Bollobás and F. R. K. Chung. The diameter of a cycle plus a random matching. *SIAM J. Discret. Math.*, 1(3):328–333, 1988.

[4] B. Bollobás and W. F. de la Vega. The diameter of random regular graphs. *Combinatorica*, 2(2):125–134, 1982.

[5] C. Cattuto, C. Schmitz, A. Baldassarri, V. D. P. Servedio, V. Loreto, A. Hotho, M. Grahl, and G. Stumme. Network properties of folksonomies. *AI Commun.*, 20(4):245–262, 2007.

[6] E. H. Chi and T. Mytkowicz. Understanding the efficiency of social tagging systems using information theory. In *HT '08: Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*, pages 81–88, New York, NY, USA, 2008. ACM.

[7] A. Dieberger. Supporting social navigation on the world wide web. *Int. J. Hum.-Comput. Stud.*, 46(6):805–825, 1997.

[8] T. Eda, T. Uchiyama, T. Uchiyama, and M. Yoshikawa. Signaling emotion in tagclouds. In *WWW '09: Proceedings of the 18th international conference on World wide web*, pages 1199–1200, New York, NY, USA, 2009. ACM.

[9] P. Erdos and A. Renyi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5:17–61, 1960.

[10] W.-T. Fu and P. Pirolli. Snif-act: a cognitive model of user navigation on the world wide web. *Hum.-Comput. Interact.*, 22(4):355–412, 2007.

[11] M. A. Hearst and D. Rosner. Tag clouds: Data analysis tool or social signaller? In *HICSS '08: Proceedings of the Proceedings of the 41st Annual Hawaii International Conference on System Sciences*, Washington, DC, USA, 2008. IEEE Computer Society.

[12] D. Helic, C. Trattner, M. Strohmaier, and K. Andrews. On the navigability of social tagging systems. In *Proc. of 2010 IEEE International Conference on Social Computing*, pages 161–168, Los Alamitos, CA, USA, 2010. IEEE Computer Society.

[13] P. Heymann, A. Paepcke, and H. Garcia-Molina. Tagging human knowledge. In *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, pages 51–61, New York, NY, USA, 2010. ACM.

[14] A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. Bibsonomy: A social bookmark and publication sharing system. In *Proceedings of the Conceptual Structures Tool Interoperability Workshop at the 14th International Conference on Conceptual Structures*, pages 87–102, 2006.

[15] O. Kaser and D. Lemire. Tag-Cloud Drawing: Algorithms for Cloud Visualization. *Proceedings of Tagging and Metadata for Social Information Organization (WWW 2007)*, 2007.

[16] J. Kleinberg. The small-world phenomenon: An algorithmic perspective. In *Proceedings of the 32nd ACM Symposium on Theory of Computing*, 2000.

[17] J. M. Kleinberg. Navigation in a small world. *Nature*, 406(6798), August 2000.

[18] J. M. Kleinberg. Small-world phenomena and the dynamics of information. In *Advances in Neural Information Processing Systems (NIPS) 14*, page 2001. MIT Press, 2001.

[19] J. Kleinfeld. Could it be a big world after all? The six degrees of separation myth. *Society*, April 2002.

[20] M. Kochen, editor. *The Small World*. Ablex, Norwood, NJ, 1989.

[21] C. Körner, R. Kern, H. P. Grahsl, and M. Strohmaier. Of categorizers and describers: An evaluation of quantitative measures for tagging motivation. In *21st ACM SIGWEB Conference on Hypertext and Hypermedia (HT 2010), Toronto, Canada, ACM*, Toronto, Canada, June 2010.

[22] J. Leskovec, J. Kleinberg, and C. Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. In *KDD*

'05: Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining, pages 177–187, New York, NY, USA, 2005. ACM.

[23] R. Li, S. Bao, Y. Yu, B. Fei, and Z. Su. Towards effective browsing of large scale social annotations. *Proceedings of the 16th international conference on World Wide Web*, page 952, 2007.

[24] C. S. Mesnage and M. J. Carman. Tag navigation. In *SoSEA '09: Proceedings of the 2nd international workshop on Social software engineering and applications*, pages 29–32, New York, NY, USA, 2009. ACM.

[25] P. Mika. Ontologies are us: A unified model of social networks and semantics. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(1):5–15, 2007.

[26] S. Milgram. The small world problem. *Psychology Today*, 1:60–67, 1967.

[27] D. Millen and J. Feinberg. Using social tagging to improve social navigation. In *Workshop on the Social Navigation and Community Based Adaptation Technologies. Dublin, Ireland.* Citeseer, 2006.

[28] N. Neubauer and K. Obermayer. Hyperincident connected components of tagging networks. In *HT '09: Proceedings of the 20th ACM conference on Hypertext and hypermedia*, pages 229–238, New York, NY, USA, 2009. ACM.

[29] P. Pirolli. An elementary social information foraging model. In *Proceedings of the 27th international conference on Human factors in computing systems*, pages 605–614. ACM, 2009.

[30] M. Ramezani, J. Sandvig, T. Schimoler, J. Gemmell, B. Mobasher, and R. Burke. Evaluating the impact of attacks in collaborative tagging environments. In *Computational Science and Engineering, 2009. CSE '09. International Conference on*, volume 4, pages 136 –143, aug. 2009.

[31] A. W. Rivadeneira, D. M. Gruen, M. J. Muller, and D. R. Millen. Getting our head in the clouds: toward evaluation studies of tagclouds. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 995–998, New York, NY, USA, 2007. ACM.

[32] C. Schmitz, A. Hotho, R. Jäschke, and G. Stumme. Mining association rules in folksonomies. In *Data Science and Classification: Proc. of the 10th IFCS Conf., Studies in Classification, Data Analysis, and Knowledge Organization*, pages 261–270. Springer, 2006.

[33] J. Sinclair and M. Cardew-Hall. The folksonomy tag cloud: when is it useful? *Journal of Information Science*, 34:15, 2008.

[34] M. Strohmaier, C. Koerner, and R. Kern. Why do users tag? Detecting users' motivation for tagging in social tagging systems. In *International AAAI Conference on Weblogs and Social Media (ICWSM2010), Washington, DC, USA, May 23-26*, 2010.

[35] M. Strohmaier, C. Trattner, and D. Helic. The benefits and limitations of tag clouds as a tool for social navigation from a network-theoretic perspective. *Journal of Universal Computer Science*, 2010. (UNDER REVIEW).

[36] D. J. Watts, P. S. Dodds, and M. E. J. Newman. Identity and search in social networks. *Science*, 296:1302–1305, 2002.

[37] D. J. Watts and S. H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393(6684):440–442, June 1998.

[38] Y. Zhang, B. Jansen, and A. Spink. Time series analysis of a Web search engine transaction log. *Information Processing & Management*, 45(2):230–245, 2009.

4

Evaluating tag-based Information Access in Image Collections

This chapter is based on the paper *"Evaluating Tag-Based Information Access in Image Collections"* which was presented at the 23rd ACM Conference on Hypertext and Social Media in 2012.

It continues the work on the usefulness of tags to support efficient search and navigation in tagging systems. In particular, this chapter presents a controlled user study that compares three tag-based search interfaces on two recognized types of search tasks – lookup and exploratory search – with each other. The interfaces explored in the study include a standard search interface that plays the role of a baseline and two types of tag-based search interfaces: a regular search interface using traditional tag clouds and a faceted interface using classified tags. We demonstrate that tag-based search interfaces significantly outperform traditional search-only interfaces in both performance and user satisfaction.

The original contribution was published in the proceedings of the conference and can be found in [37].

4.1 Abstract

The availability of social tags has greatly enhanced both search-based and browsing-based access to information. Tag clouds, emerged as a new "social" way to find and visualize information providing both one-click access to information and a snapshot of the "aboutness" of a tagged collection. A range of research projects explored and compared different types of tag artifacts for information access ranging from regular tag clouds to tag hierarchies. At the same time, there is a lack of user studies that compare the efficiency of different types of tag-based browsing interfaces from the users point of view. This paper contributes to the research on tag-based information access by

presenting a controlled user study that compared three types of tag-based information access interfaces on two recognized types of search tasks – lookup and exploratory search. Our results demonstrate that tag-based browsing interfaces significantly outperforms traditional search-only interfaces in both performance and user satisfaction. At the same time, the differences between the two types of tag-based browsing interfaces explored in our study are not as clear.

4.2 Introduction

Social tags provide an easy and intuitive way to annotate, organize and retrieve resources from the Web. Promoted by several pioneering systems such as Delicious, Flickr, and CiteULike, social tagging has emerged as one of the most popular technologies of the modern Web. The value of tags was specifically advocated for image collections such as Flickr where the presence of tags made images searchable and discoverable. While tags help to discover content even with a standard keyword-search, the most innovative feature of social tags was the ability to support browsing-based access to information through so-called "tag clouds". Effectively, tag clouds, are a new "social" way to find and visualize information providing both: one-click access to information and a snapshot of the "aboutness" of a tagged collection. Not surprisingly, a large volume of research was devoted to developing better approaches to construct and visualize tag clouds [5, 31, 18] as well as more advanced tag constructs such as clustered/classified tag clouds [23, 33, 2, 41, 16, 25] and tag hierarchies [10, 19, 35, 36].

The majority of research on tag clouds and hierarchies used an information- or network-theoretical approach to evaluate the quality of different tag constructs by terms of search and navigation and ignores the user prospective. User studies comparing performance of users applying different tag-based browsing constructs in a set of realistic search tasks are rare. Moreover, there is a lack of user studies that compare the effectiveness of various tag constructs against simple search-based access to tagged collections. This paper attempts to bridge this gap by comparing several types of tag-based information access in a controlled user study. The study has been performed in the context of image search where the presence of tags is known to be most valuable. To make the study more useful, we compared the performance of three types of tag-based information access interfaces in two commonly recognized types of search tasks – lookup search and exploratory search. The tag-based interfaces explored in the study include a search-based interface that plays the role of a baseline and two types of tag-based browsing interfaces: a regular browsing interface using traditional tag clouds and a faceted browsing interface using classified tag clouds. We selected the faceted tag cloud interface from among other advanced tag-based browsing approaches because our previous study [26] in the image search domain revealed that faceted search interfaces helped users to better explore large collections of images.

4.3 DataSet

As dataset for our study we utilized a collection of images from an archive belonging to the Carnegie Museum of Art in Pittsburgh, Pennsylvania. Overall, the collection contains more than 80,000 images taken by the famous local photographer Charles Teenie Harris, who captured African-American life in Pittsburgh over a 40-year period. In our study, we used 1,986 of these images, of which 986 have been featured in a current exhibition at the Carnegie Museum of Art. The remaining 1000 images were included in this study as they provide a good overview of the entire collection and represented well in corresponding exhibition categories. For the 1,986 images, we collected user tags using the Amazon Mechanical Turk[1]. Overall, the dataset provides 4,206 unique tags and 16,659 tag assignments applied by 97 users for the 1,986 images.

4.4 Interfaces

For the purpose of our study, we implemented three tag-based interfaces to search the collection of Teenie Harris images – one standard "search box" interface and two interfaces that support both search and tag-based browsing. In the following section, we introduce these interfaces and their functionalities.

4.4.1 The Baseline (Search Only) Interface

As a baseline for our study (see Figure 4.1), we utilized a simple search box-based interface that offers the look and feel of well-known search engines. Similar to the Google, Yahoo! or Bing image search interfaces, we provide our users with a search box to issue a query, a thumbnail preview of the resulting images sorted by relevance and the functionality to click on the image in order to get a more detailed view of the image resource. The back-end of our search interface is built upon the OpenSource search engine Apache Lucene[2], which utilizes the tags of each image to create the search index.

[1] https://www.mturk.com/

[2] http://lucene.apache.org/java/docs/index.html

Figure 4.1: Screenshot of the baseline interface.

4.4.2 The Tag Cloud Interface

The second interface explored in this paper is referred to as the tag cloud interface. As indicated by its name, this type of search interface extends the baseline search interface with the functionality of a traditional tag cloud. The alphabetically ordered tag cloud provided the user with a topical overview of the search results and allow the user to search or browse images using the tags displayed in the cloud. This form of tag cloud is currently the most popular type of tag-based browsing in social tagging systems. To generate the tag cloud in this interface, we utilized a simple popularity based tag cloud algorithm. For each query, we display the top N most frequent co-occuring tags to the user. This approach was shown to be the one of the best choices to create a tag cloud from the prospective of tag-based search and browsing [39]. It is currently the most popular algorithm to generate tag clouds. Since the number of tags displayed in the tag cloud is an important factor which was shown to negatively affect tag cloud-based search and navigation [34, 20], we also provide the functionality to increase or decrease the number of tags in the tag cloud to suit the user's needs. In Figure 4.2, a sample screenshot is presented to show how the tag cloud interface appears on the user's screen. As can be seen in the figure, the interface offers not only the functionality to click on a tag to issue a query, but also the possibility to expand the query

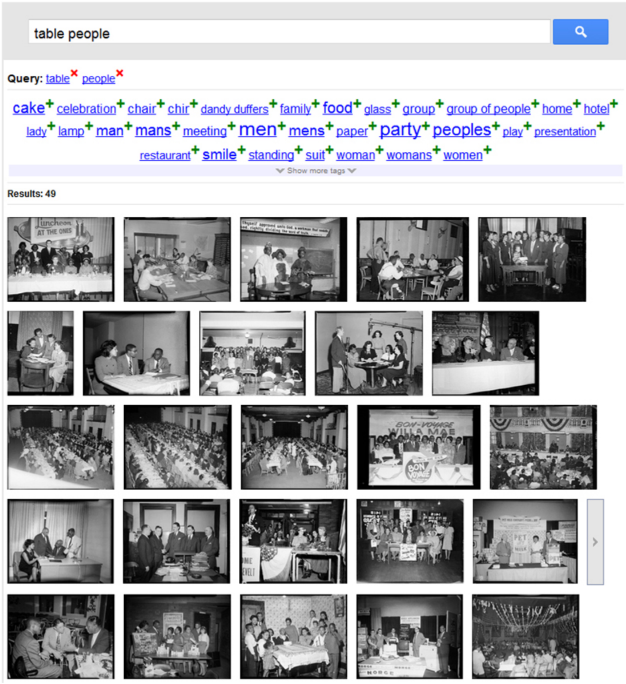


Figure 4.2: Screenshot of the tag cloud interface.

by clicking the "+" sign in the tag cloud or shrink the query by utilizing the "x" sign in the query string beneath the search box. Currently, many popular tagging systems such as Delicious or BibSonomy offer similar approaches for query expansion or reduction to support the user with a more flexible way to search and navigate in a tag based information system.

4.4.3 The Faceted Tag Cloud Interface

The third interface developed for the study is referred to as a faceted tag cloud interface (see Figure 4.3). It can be considered as one of the most innovative tag-based search interfaces currently available. The interface was first introduced in 2009 by Yahoo! [33] in order to search for images in the social tagging system Flickr. Although there are very few implementations of this type of interface, there is a great deal of current research in this area [29, 40, 8, 7]. Similar to the tag cloud interface, this type of interface provides the user with the functionality to view the tags of the retrieved images in a visually appealing representation. However, contrary to the traditional tag cloud interface, where all tags appear in a tag cloud in an unstructured way, this type of interface classifies tags into several categories.

To decide which classification schema to utilize, we performed an extensive literature survey on currently available tag classification approaches

[6, 29, 40, 8, 33, 11]. In the end, we selected a simplified form of the well known "Editor's 5 Ws" approach that recognizes "Who" (people, groups or individuals), "Where" (location or places), "When" (time, activities or events), "What" (objects, food, animals or plants) and "Other" (unknown, not classified) classification schema. This schema was found to be effective in classifying tags in the image domain [33] as well as in our earlier user studies [26]. To classify our tags for this type of interface, we also used Amazon Mechanical Turk. The classification procedure itself was independent of image context as none of the currently available tag classification approaches take into account context information such as resource information, user information or other tags for the same or similar resources.

To ensure that the workers on Amazon Mechanical Turk (referred to as turkers) would classify our tags in a meaningful way, we provided them detailed instructions of how to select those tags which fit into the one of the five given categories. The guidance included a sample screenshot of three different types of tags classified into one of the five categories and a detailed explanation of how to use these categories. Overall, three turkers were assigned to classify each particular tag. After the first classification round, we noted that 11% of tags were not classified as the turkers could not agree on which of the five given categories to use. Therefore, we decided to initiate a second classification round with an additional six turkers (per tag) to increase the precision of our classification procedure. All in all, 22% of the tags were classified as "Who", 16% as "Where", 23% as "When", 34% as "What" and only 5% of the tags as "Other", which clearly out-performs current automatic tag classification approaches by terms of not classifiable tags (represented as "Other" tags in our classification schema). We had 86 different turkers for the first classification round and 35 turkers for the second. The mean inter-rater agreement per tag over all turkers was substantial (75%).

In Figure 4.3 one can see a screenshot of how this type of interface appears on the user's screen. As with the tag cloud interface, users have the opportunity to issue a query by clicking on a tag, to expand a query by clicking on the "+" sign or shrink the query by utilizing the "x" sign in the query string beneath the search box. In addition, the faceted tag cloud can be expanded or collapsed one study session.

4.5 User Study Design

To compare the three tag-based information access interfaces, we designed a within-subject study. In this design, each of our subjects evaluated the three different search interfaces during one study session. To determine when tag-based support is most effective; each interface was examined in the context of two kinds of search tasks, which are discussed in the following section.

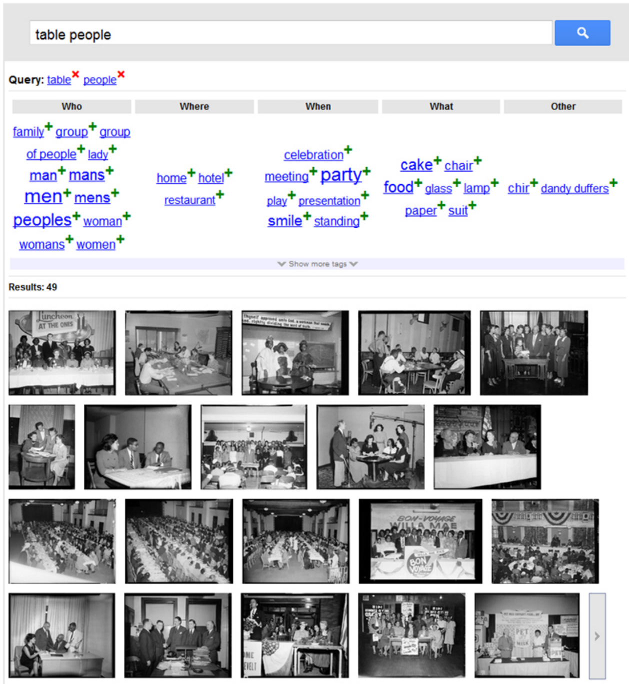


Figure 4.3: Screenshot of the faceted tag cloud interface.

4.5.1 Search Tasks

It has been shown that search task attributes affect the information seeking behavior of users [13, 38, 9]. The complexity, familiarity, clarity and difficulty of a search task influences how a person searches, browses and uses information systems [13, 17]. To account for the impact of these factors, our study separately evaluated the effectiveness of the three tag-based information access interfaces in the two primary types of search tasks known as lookup search and exploratory search.

As indicated by its name, lookup search is typically performed to find a specific information item in a document collection [27]. Lookup search tasks are considered to be relatively simple and most frequently involve using a traditional search interface (cf. [13, 38, 9]). More complicated search tasks "beyond lookup" are typically called exploratory search tasks [27, 9]. Exploratory search assumes that the user has some broader information need that cannot be simply met by a "relevant" information item (as in simple lookup search), but requires multiple searches interwoven with browsing and analysis of the retrieved information [26].

To study lookup search behavior, we created nine different lookup search tasks. All of these tasks were of similar nature: the subject was given and the user was expected to find relevant images in the collection within a certain time limit. To account for the differences in difficulty [13, 38, 9] a variety of pictures were selected ranging from "easy" to "hard" to find. To classify images by difficulty, we calculated the mean search time of each image in the image collection based on lookup searches performed on Amazon Mechanical Turk. Then, we selected nine images ranging from "easy" to "hard" to find in the Teenie Harries image collection. In Table 4.1, the nine different images chosen for the user study are presented.

To study exploratory search behavior, we designed three explor- atory search tasks as shown in Table 4.1. To ensure the balance between each type of user interface and also to capture the attribute of difficulty, we designed the exploratory search tasks carefully with a variety of additional search criteria and attributes. For instance, to capture balance with the faceted search interface, we tried to tune our search tasks to utilize as many facets as possible. We did that by asking the subjects to search for several different topics such as music, sports or shops as well as various search criteria such as different locations. To capture the property of familiarity with the search tasks, we asked our subjects in the post-questionnaire to rate their expertise level on the given topic or search item.

To be sure that our search tasks were meaningful, we performed several trial searches on Amazon Mechanical Turk and we conducted a pilot study.

Search Tasks	Search Task Descriptions
Lookup	Find the following picture!* 
Exploratory	1. Find at least 8 different types of stores/shops in Pittsburgh! Each type of store/shop should have at least two images from different locations, i.e. in total you will have to find at least 16 images.
	2. Pittsburgh is a city with many sport teams. Find at least 8 different sport activities! Each type of sport should be represented by at least two pictures. In total, you will have to get at least 16 pictures.
	3. Pittsburgh has a rich cultural heritage. There were many musicians who worked in Pittsburgh. Find at least 5 different types of music instruments which the musicians played in Pittsburgh. Each instrument needs 2 pictures and all pictures should be taken in different locations. In total, you will have to collect at least 10 pictures.

Table 4.1: Search tasks and descriptions (*= in the user study only one image was presented to the user at one time).

4.5.2 The Procedure

As discussed previously, our subjects had to undertake two different kinds of search tasks using three different types of search interfaces within one user study session. During the study, each subject was assigned to perform nine different lookup and three different exploratory search tasks which were the same for the duration of the whole experiment. To counter the impact of fatigue and learning, the order in which the search tasks and system interfaces were used were rotated using a Latin square design. In addition to this, the lookup and the exploratory search tasks were randomized among all three interfaces to make sure that each of them was evaluated under different search interface conditions. The procedure of the user study was as the follows:

1. Each participant was informed of the objective of the study and asked to complete a consent form.

2. The participant was asked to complete a short questionnaire eliciting background information.

3. For each type of interface and task, a demonstration was given. After that, the participant had plenty of time to familiarize themselves with the interfaces and tasks.

4. For each interface the user was given three lookup tasks and one exploratory search task.

    (a) For each lookup task: an image was presented to the participant and a limit of 3 minutes (+30secs. for task reading) was given to complete the task. After that, a post-search questionnaire was given to the subject to elicit disposition toward the system interface.

    (b) For each exploratory task: a description of the task was given to the participant and they were allotted a limit of 10 minutes (+1min. for task reading) to complete the task. Then, a post-search questionnaire was handed out to the subject to elicit disposition toward the system interface.

5. A final questionnaire was given to the subject to assess the differences among the three search interfaces.

6. A series of open-ended questions were asked according to the observations made during the study.

7. The participant was paid and thanked.

4.5.3 Participants

Our study involved 24 participants (8 females, 16 males), who were recruited via email and flyers distributed throughout the University of Pittsburgh campus. The participants were from a variety of disciplines ranging from law to computer science. Four of them had earned a bachelor degree, 16 a master's degree and four a PhD degree. The average age of the participants was 30.6 years (min=22, max = 61, SD=7.59 years). Almost all (except 2 participants) reported to use computers more than 5 hours a day. All participants (expect two) rated their search engine skills as high and indicated to use Google, Yahoo! or Bing frequently. A significant number (19) reported that they were familiar with tagging or use search tagging systems such as BibSonomy, Delicious or Flickr regularly. Four participants reported that they were familiar with the history of Pittsburgh, the rest of our subjects stated that they were not. On average, one user study session lasted 90 minutes.

4.6 Results

In this section we present the results of our user study. We start by comparing user performance in different search interfaces and follow with an extensive log analysis that describes how the interfaces were used. After that, we report the findings from our post and final questionnaires and report the participants subjective opinions about these interfaces.

4.6.1 Performance Analysis

The main goal of this study was to compare user search performance in two types of search tasks (lookup and exploratory search) and with three different interfaces (with and without tag-based browsing support). To assess user performance, we examined search time and number of total interface actions [24] which are traditionally used in the study of search and browsing interfaces. Shorter search time and a lower number of actions should indicate a more efficient interface for image search.

While these two performance measures are known to be reliable, they do not allow us to clearly distinguish between several search conditions in the presence of many failed search attempts (i.e., cases where the subjects were not able to complete the task and were interrupted). Due to the presence of this cap, the time and actions spent on failed attempts flattens the overall differences, making different conditions look closer than they are in reality. To avoid this problem, we separately measured user performance only on successful tasks. Given comparable success rates (as we observed in the study), user performance on successful tasks enables us to more easily distinguish between several conditions.

Table 4.2 provides a summary of performance data for our three interfaces and two kinds of search tasks. The table separately reports performance data for all tasks (including failed tasks with capped time) and only for successfully completed search tasks. As the data shows, the main difference in user performance is observed between the task types: exploratory search, as expected, required much more time and actions than lookup tasks. To discover significant performance differences among interfaces, we applied 2 x 3 ANOVA (analysis of variance). The analysis was done separately for search time and for the total number of interface actions as functions of search task and interface. We also separately evaluated data for all cases and for successful cases only. The analysis of successful cases data revealed significant differences between tag cloud and baseline interfaces in terms of search time, $p < .001$, and total actions, $p < .001$, under exploratory search. Likewise, we found a significant difference on the number of total interface interactions between faceted tag cloud and baseline(search only), $p = .037$. No significant differences were discovered for "the data for all cases". We also have not discovered any significant differences between the two kinds of tag-based browsing interfaces under all conditions.

Effect of familiarity and difficulty on performance. Prior research on exploratory search interfaces indicated that the value of advanced information access interfaces might depend not only on the type of task (i.e., lookup vs. exploratory search) but also on task difficulty [13] and user familiarity with the search topic [17]. In the context of our study, we registered some reasonable differences in user familiarity on a Liker scale(1-5) with the topics of the three exploratory search tasks (M=*3.125*, SE=.15056, SD=1.27751). In other words, it was possible to divide users into two groups for each task - those familiar with the task topic and those not. Moreover, as the study indicated, the level of difficulty in the three exploratory search tasks was considerably different between the one relatively easy task and the two more complicated tasks. These variations allowed us to perform a separate analysis that explored the combined effect of the interface, task difficulty, and task familiarity in the context of exploratory search. We ran a 3 x 3 ANOVA as a function of task difficulty and interface, and also controlling for the two levels of familiarity previously mentioned. As shown in Table 4.3, the analysis revealed a significant difference between tag cloud and baseline interfaces in search time for those users not familiar with the topic and at a medium level of task difficulty when considering all cases, $p = .014$, and when only considering successful cases, $p = .009$. No other significant differences were found. These results indicate that the tag cloud interface provides the most significant impact in cases where tasks are more complicated and users are less familiar with the topic of the task.

A similar analysis of the impact of difficulty and familiarity was performed for the lookup search context, but we did not find significant differences between interfaces. However, the impact of difficulty and familiarity might be determined by the relatively low comparable level of user task familiarity in this context. On average of the ratings in the lookup search task (M=1.3611, SE=.08463, SD=.71809), our subjects were not as familiar with the images as they were in the exploratory task of the user study. Only two of them reported that they were familiar with the images due to the fact they found an image during the search session of a previous task.

Task	Measure	Baseline		Tag Cloud		Facet	
		All cases	Successful	All cases	Successful	All cases	Successful
Lookup	Cases	72	59	72	57	72	59
	Total Actions	9.01±.89	6.46±.67	8.58±.94	5.37±.56	8.68±.86	6.12±.63
	Search Time	77.35±7.35	54.19±5.31	75.38±8.03	44.37±4.48	77.67±7.8	52.17±5.32
Exploratory	Cases	24	23	24	20	24	22
	Total Actions	43.67±4.36	**42.17±4.27**	41.04±4.52	**33.50±3.37****	42.58±4.26	40.73±4.44
	Search Time	421.58±38.03	**413.48±38.81**	363.96±35.05	**312.4±30.74*****	378.33±33.46	356.91±32.8

Table 4.2: Descriptives of total actions and search time by search and interface. Each statistic is calculated considering all cases and considering only successful search tasks (**=significant at $p<0.01$; ***=significant at $p<0.001$).

Difficulty	Measure	Baseline		Tag Cloud		Facet	
		All cases	Successful	All cases	Successful	All cases	Successful
Hard	cases	6	6	7	4	6	4
	Total Actions	67.33±5.94	67.33±5.94	64.43±8.48	51.5±10.9	55.5±6.59	51.75±9.71
	Search Time	603.5±23.05	603.5±23.05	557.43±40.42	507.5±61.4	562.67±38.47	537.0±55.14
Medium	cases	3	3	4	4	3	3
	Total Actions	38.33±5.24	38.33±5.24	35.25±3.09	35.25±3.09	57.33±6.89	57.33±6.89
	Search Time	**494.67±148.17***	**494.67±148.17****	**285.75±16.95**	**285.75±16.95**	382.00±22.11	382.00±22.11
Easy	cases	5	5	5	5	6	6
	Total Actions	25.0±4.24	25.0±4.24	23.6±2.5	23.6±2.5	19.0±1.53	19.0±1.53
	Search Time	308.8±49.31	308.8±49.31	227.8±23.77	227.8±23.77	212.23±25.45	212.33±25.45

Table 4.3: Descriptives of total actions and search time, $mean \pm SE$, by interface at different difficulty levels, when people are not familiar with the topics and under exploratory search tasks (*=significant at $p<0.05$).

4.6.2 Looking Deeper: Log Analysis

Although the previous analysis reveals performance differences between interfaces and tasks, it does not show how different usage profiles were for each of the interfaces and tasks. To look for these differences we performed extensive user log analysis on answering specific questions.

The first question was *How different were usage profiles for different interfaces and tasks?* To build the usage profile, we distinguished several different interface actions: (1) *Search* (inserting a query in the search box); (2) *Click Tag* (issuing a query by clicking on a tag); (3) *Add Tag* (expanding the query with a tag by clicking the "+" sign); (4) *Remove Term* (removing a term from the query by clicking the "x" sign); (5) *Show More Tags* (clicking the show more tags button to increase the number of tags in the tag cloud): (6) *Show Fewer Tags* (clicking the show fewer tags button to reduce the number of tags in the tag cloud); (7) *Show More Results* (clicking the show more results button to increase the number of images in the result list); (8) *Click Image* (clicking on an specific image) and (9) *Total Actions*.

Table 4.5 presents usage profiles for different interfaces and search tasks. The most visible (albeit trivial) result is that the action *Search* is used more frequently in the baseline interface, $p = .006$. While the *Search* action is also used more frequently in the tag cloud than in the faceted tag cloud interface, this difference is not significant. Another interesting discovery is that the use of *Show More Results* is significantly higher in the baseline interface than in the tag cloud, $p = .015$. The corresponding difference between the baseline and the faceted tag cloud is close to significant at the acceptable level $p = .055$. Since the use of *Show More Results* is the evidence that the top results returned by the last search or tag browsing action were not satisfactory, we can argue that tag browsing was more successful at providing relevant results. We can speculate that this result stems from the tag browsing interface's ability to provide a snapshot of the "aboutness" of the collection, guiding the user to a more successful choice of a search term or tag. In addition, we found an intriguing difference between the tag cloud and the faceted tag cloud interfaces: the action *Add Tag*, which was used to narrow the results by adding tags to the query, was used significantly more frequently in the faceted interface than in the tag cloud interface, $p = .006$. The difference among interfaces in terms of the usage frequency of other actions (*Click Tag*, *Remove Term*, *Show More Tags*, *Show Less Tags*) was not significant. Table 4.5 also reports differences in the usage profile between lookup and exploratory search tasks. As we can see, the usage profile was considerably different for the two types of tasks. This emphasizes that lookup and exploratory search tasks are radically different from the user perspective. However, as users had different amount of time available to complete lookup and exploratory search tasks, we compared percentages instead of the mean number of actions. However, to test for significant differences between these

	Who	Where	When	What	Other
Baseline	9.9%	29.6%	11.7%	42.7%	6.2%
Tag Cloud	13%	28.8%	9.4%	43.2%	5.6%
Facet	16.2%	24%	18.9%	34.6%	6.3%

Table 4.4: Percentage of search actions in each type of semantic category by search interface.

percentages, we run one chi-squared test per each action. As shown in Table 4.5, we found significant differences for the *Search* action, $p < .001$, the *Add Tag* action $p < .001$, the *Remove Term* Action, $p < .001$ and the *Show More Results* action $p < .001$. These indicate that people rely more on the search box, the *Add Tag* and *Remove Term* functionality, and skimming through the paginated result list in lookup tasks than in exploratory search tasks. The significant difference on *Click Image* action, $p < .001$, shows that people rely more on clicking images in exploratory search than in lookup search.

The second question that we attempted to answer was *Does tag grouping by semantic category affect the usage of these categories?* As outlined in Section 4.4.3, we classified tags in our tag corpus into the following five dimensions: Who, Where, When, What and Other. The users in the faceted interface case were able to see which tag belongs to which category. However, the users of both the search and regular tag cloud interfaces used the same terms in search and browsing, although without knowing to which category the issued query term or the clicked tag belonged. One could hypothesize that the tag usage profile (i.e., frequencies of using tags in different categories) may be affected by making the categories visible. Table 4.4 shows the proportion of query terms in each classification category as used by the study participants; each row presents percentages in each type of interface. We analyzed the significant difference in these percentages by running two chi-square goodness of fit tests. Considering overall tag usage,(i.e., aggregating lookup and exploratory search tasks), as well as setting the expected percentages of the tag categories to match those in the faceted tag cloud interface, we found them significantly different than those in the baseline interface ($\chi^2(4,548) = 46.092$, $p < .001$), and the percentages in the tag cloud interface ($\chi^2(4,683) = 58.612$, $p < .001$). In particular, we see a visible increase in *Who* tags at the expense of *What* tags. This data provides evidence that explicit tag categorization does impact user behavior.

	Interface			Task			
	Baseline	Tag Cloud	Facet	Lookup	%	Exploratory.	%
Search	**9.24±.96****	**5.61±.82**	**4.81±.63**	3.89±.28	**44.45%*****	14.54±1.36	**34.27%**
Click Tag	.00	2.88±.46	2.92±.46	.94±.13	10.67%	4.92±.74	11.58%
Add Tag	.00	**.61±.14**	**1.25±.2****	.72±.11	**8.19%*****	.33±.12	**0.78%**
Remove Term	.00	.95±.18	1.40±.25	.62±.1	**7.08%*****	1.26±.3	**2.97%**
Show More Tags	.00	.17±.07	.11±.05	.065±.02	0.74%	.18±.09	0.42%
Show Less Tags	.00	.02±.01	.00	.01±.0	0.05%	.01±.01	0.03%
Show More Results	**1.78±.3****	**.86±.18**	1.01±.19	1.31±.17	**14.90%*****	.96±.2	**2.25%**
Click Image	6.66±1.1	5.59±.86	5.66±.87	1.22±.07	**13.90%**	20.22±.99	**47.65%*****
Total Actions	17.68±1.99	16.70±1.95	17.16±1.94	8.76±.52	100%	42.73±2.5	100%

Table 4.5: Summary of the means of actions based on each task session in the baseline, tag cloud, faceted tag cloud interfaces and means/percentages of actions based on each task session and interface for lookup and exploratory search tasks (**=significant at p<0.01, ***=significant at p<0.001).

4.6.3 Post-Task Questionnaires: Participants' Perceptions of the Interfaces

To better understand the participants' perceptions of each interface, we focus on analyzing user feedback about the different interfaces and their features. In the user study, the participants were asked to compete a post-task questionnaire after each of their search tasks was finished. By analyzing this questionnaire, we could assess the usefulness of each interface and see whether any significant differences could be found among the three interfaces and also between two search tasks (lookup vs. exploratory). Table 4.6 shows the average user rating for each question in the survey.

In Question 1 and 2, a 2 x 3 ANOVA was conducted on users' ratings in order to examine the effect of interface and search task. There is no significant interaction between interface and search task. For Question 1, a simple main effect analysis showed that there is a significant difference between the interfaces $F(2,46) = 30.113$, $p < .001$. Participants judged the support provided by the tag cloud interface significantly better than that provided by the baseline, $p < .001$. They also rated the interface support of the faceted tag cloud interface significantly better than that of the baseline, $p < .001$.

For Question 2, we also found a significant difference between the interfaces $F(1.406,32.332) = 11.097$, $p = .001$. Participants felt that the baseline interface had less "unnecessary features" than tag cloud, $p < .001$, and the faceted tag cloud, $p < .001$. However, the unnecessary features were a relatively trivial concern to the users of all three interfaces.

Question 3 specially asked about the exploratory search task, "How confident were the participants on the systems' ability to find relevant information". A 1-way ANOVA was used to test for performance differences among the three interfaces. We found a significant difference among the interfaces $F(2,46) = 5.412$, $p = .008$. The participants were significantly more confident in the ability to find relevant information with the tag cloud interface, $p = .015$, and the faceted tag cloud interface, $p = .037$, compared to the baseline interface.

In Questions 4–7, we investigated the usefulness of various tag-related features. The 2 x 2 ANOVA as a function of interface (tag cloud and faceted tag cloud interfaces) and search task showed that the only significant difference within this group of questions "Was the x helpful to remove terms from the query", $F(1,20) = 6.450$, $p = .02$. The result indicated that users found this interface feature significantly more useful in the tag cloud that in the faceted tag cloud interface. No significant difference was found between the lookup and the exploratory search tasks in respect to Question 8.

Question	Lookup Task			Exploratory Task		
	Baseline	Tag Cloud	Facet	Baseline	Tag Cloud	Facet
1. Did the interface provide enough support for that task?	**2.88±.24**	**3.92±.15***	**4.04±.15****	**2.88±.21**	**4.21±.13***	**4.21±.13***
2. Were some of the interface features unnecessary for that task?	**1.33±.12**	**1.83±.18***	**1.92±.2***	**1.33±.12**	**1.54±.13***	**2.17±.23***
3. Were you confident in the system's ability to find relevant information on this topic?	-	-	-	**3.25±.22**	**3.92±.2****	**3.92±.18****
4. Did you find the tag cloud/faceted tag cloud helpful in finding relevant information?	-	3.79±.2	3.96±.18	-	4.13±.22	3.83±.21
5. Was it helpful to display the tags in different font sizes?	-	3.5±.23	3.54±.23	-	3.17±.27	3.38±.24
6. Was the + useful to add terms to the query?	-	3.77±.25	3.82±.27	-	4.05±.2	3.73±.27
7. Was the x helpful to remove terms from the query?	-	**4.09±.23****	**3.65±.26**	-	4.04±.17	4.04±.18
8. Did you find it distracting that some terms in the faceted tag cloud were not classified correctly?	-	-	2.33±.26	-	-	2.43±.25

Table 4.6: Mean average of response to post questionnaire items (*=significant at p<0.05; **=significant at p<0.01, scale 1-5, higher values indicate more agreement with the statement).

Question	Interface Baseline (f.)	 Tag Cloud (f.)	 Facet (f.)
1. Which of the interfaces did you prefer most?	4.2% (1)	**54.2% (13)**	41.7% (10)
2. Which of the interfaces would you prefer for lookup search?	4.2% (1)	41.7% (10)	**54.2% (13)**
3. Which of the interfaces would you prefer for exploratory search?	- (-)	41.% (10)	**58.3% (14)**
4. Which of the interfaces would you suggest the Carnegie Museum of Art?	- (-)	41.% (10)	**58.3% (14)**

Table 4.7: Percentages and frequencies (=f.) to final questionnaire items.

4.6.4 Post Questionnaires: Participants' Interface Preferences and Comments

Another useful source of user feedback was a post questionnaire that was administered after each participant completed the entire study. This questionnaire offered us an opportunity to ask users about their opinions about three different interfaces. By this point in time, users had gain practical experience with both types of tasks and all three types of interfaces. As shown in Table 4.7, when asked a retrospective question "*Which one of the interfaces did you like/prefer most?*", 54.2% (13) of subjects preferred the tag cloud interface, 41.7% (10) the faceted tag cloud interface, and only 4.2% (1) preferred the baseline search interface. This data correlates well with the users' actual performance on tasks. At the same time, user feedback differed on "forward looking" questions designed to assess user preferences in future situations (such as, "*Which one of the interfaces would you prefer for lookup search?*"). For both tasks, the faceted tag cloud interface emerged as most preferred for future use. In addition, none of the users clearly preferred the baseline interface for exploratory search tasks. It is interesting that our subjects reported divergent results when they were asked about preference in general and for each specific task.

Further, we found that 14 subjects favored the same interface for both past and future use while the other 10 subjects indicated a preference for a different interface when working on at least one type of tasks in the future. In particular, among the 10 subjects who reported changing preferences, one subject who favored the baseline (search only) interface in the prior tasks switched to the tag cloud interface for exploratory search tasks.

We believe that the most likely explanation for the difference in interface preferences between past and future tasks is the interface complexity. While the baseline search interface is very familiar to our subjects, both the tag-

Question	Rating
1. Overall how would you rate the Search interface?	**2.75±.22**
2. Overall how would you rate the Tag Cloud interface?	**4.17±.13***
3. Overall how would you rate the Faceted Tag Cloud interface?	**4.04±.15***

Table 4.8: Mean±*SE* of response to final questionnaire items (*=significant at p<0.05; higher values indicate more agreement with the statement).

based browsing interfaces were rather novel. Moreover, while the subjects might have had at least some experience with using the traditional tag cloud interface, the faceted tag cloud was new to all of them. It is reasonable that a user's opinion of a more complex interface might be less favorable during their first attempts in using it. At the same time, armed with some experience, the users expressed stronger preferences for the use of more complex and powerful interfaces in the future. This might explain the difference in users' answers to the question "*which of the interfaces they would recommend for Carnegie Museum of Art*" (i.e., to professionals working with images): 58.3% (14) of our subjects recommended the faceted tag cloud interface while only 41.7% (10) subjects recommended the tag cloud interface; no one recommended the baseline interface. This indicates that tag-based browsing interfaces, particularly the faceted tag cloud interface, were evaluated to be more powerful and more preferred for experienced users.

The data also showed that the main difference in users' perceptions is between the baseline and the two tag-based browsing interfaces. One or another tag interface was preferred almost unanimously for both previous and future situations. At the same time, the difference between the two tag-based browsing interfaces is much less pronounced: the traditional tag cloud interface appeared to be a bit simpler and more preferred during prior tasks (which correlates well with the performance data), while the faceted tag cloud was perceived as a bit more powerful and preferred for future tasks.

Further support for this assessment of users' subjective preferences across the three interfaces is provided by analyzing their explicit the rating for each interface (see Table 4.8). On a Likert scale(1-5), the average rating for the baseline (search only) interface was 2.75, 4.17 for the tag cloud interface and 4.04 for faceted tag cloud interface. From these statistics, we can see that the baseline interface was rated significantly lower than the tag cloud interface, $p = .002$, and the faceted tag cloud interface, $p < .001$. However, there is no significant difference between the tag cloud and the faceted tag cloud interfaces.

4.6.5 Looking Deeper: Comment Analysis

To explain differences in users' perceptions of the different interfaces and their features, we examined free-form comments provided in the post questionnaire. Below, these comments are grouped by the type of the interface preferred by the user:

Preferred Baseline (Search Only) Interface

According to the 24 participants, only 1 subject preferred the baseline search interface. The reason why the user chose this type of interface favorite was the following:

> "I liked the search box most, because everything else distracted me. For me it is not necessary to have tags, because I have everything in my mind!" – P20

This subject identified that the simplest interface is the best as it did not distract by adding elements to the interface.

Preferred Tag Cloud Interface

Thirteen subjects preferred the tag cloud interface. Based on the feedback from the interview and open-ended question on why they preferred a particular interface, our subjects attributed their preference of the tag cloud interface to more usefulness than the baseline interface. They also felt that it was easier to use than the faceted tag cloud interface.

> "The tag cloud provided more information than search only and was less complex than the facet search interface" – P4

> "I think the tag cloud interface was very helpful for exploratory search task and the faceted tags are a bit harder because I have to figure out what facet to look at" – P3

> "The tag cloud shows a good overview. I click more on general terms" – P10

> "I like tag cloud because it gives me new ideas and it is easier to use" – P21

Sometimes, the poor categorization of tags in the faceted tag cloud interface accounted for why our subjects preferred the non-faceted interface. They either thought the category of facet was of low quality or irrelevant to the task.

> "The facet did not seem to identify tags well" – P1

> "I would recommend the faceted interface only if tags are rich enough and categorized correctly, otherwise tag cloud is better" – P8

> "I think the categorization was not good, it was not relevant to the task" – P19

Some of the subjects preferred the tag cloud interface because they thought that the different font sizes in the tag clouds made more sense than the categorizing tags. Furthermore, some of them even didn't pay attention to the category at all.

> "I did not look at the facets at all as I just looked at the terms" – P12

> "I could see suggestions from the tag cloud but I did not pay attention on the grouping of the tags" – P17

> "Font size attracted my attention more than the facets" – P18

> "The font size helped me to get most relevant information quickly" – P24

Preferred Faceted Tag Cloud Interface

Overall, we had 10 subjects who preferred the faceted tag cloud interface. The reason for this preference can be categorized into three aspects. First, they thought that the faceted tag cloud interface provided them with more functionality.

> "I like faceted tag cloud because the interface provided me with the most functionality" – P6

> "For difficult search task the facet is useful and for easy tasks you can just ignore the facet feature" – P7

> "The Faceted tag cloud interface seems to be a smarter interface" – P13

> "I like faceted tag cloud interface the most, because the it provides me with more information" – P23

Second, our subjects opined that the faceted tag cloud interface organized tags in more meaningful ways than the tag cloud interface.

> "I prefer faceted tag cloud interface because it shows more tags in an organized way, so I could find more information faster" – P2

> "It is easy to find the tags that I needed in faceted tag cloud" – P11

> "I like faceted tag cloud interface, because the interface is clearer and I always know where to find the tag" – P15

The third aspect is that some of our subjects thought that the faceted tag cloud suggested more keywords to them. The interface also inspired them to think of additional relevant key terms.

> "I like the faceted tag cloud because it suggest more query options than the tag cloud" – P5

> "The faceted tag cloud made me think of more useful keywords than the tag cloud" – P21

4.7 Related Work

Tagging systems such as Delicious, Flickr, and CiteULike have emerged as one of the most popular technologies of the modern Web era. Tagging behavior has been widely studied with regards to either the structure of tagging systems [15, 32], or qualitative insights about tagging behaviors across small collections [3, 12, 28]. The collective tagging behavior of users seems to offer a strong platform for summarizing and indicating content popularity to improve Web search [1].

In the computer-supported cooperative work (CSCW) domain, researchers have noted that tags could be utilized to offer search signals to others in the community. Several ranking algorithms have been investigated to improve search performance within the tagging space, such as SocialSimilarityRank [4], and FolkRank [21]. In the HCI community, Furnas et al. discovered the similarities in the cognitive processes between generation of search keywords and tags [14]. Kammerer et al. investigated how to apply relevance feedback about tags to indicate users' interests in various topics as well as to enable rapid exploration of the topic space [22]. Although CSCW and HCI both have provided different approaches to improve Web search, the focus of those studies was only on optimizing search ranking algorithms.

To understand how people use tags in reality and to what extent tag-based browsing constructs support users during their information seeking processes, we are interested in exploring the usage and efficiency of tag-based search interfaces. From an interface point of view, several interfaces have been explored. While tags are used to discover content in a traditional keyword-based search context, the innovative usage of social tags also supports browsing-based access to information. For instance, in [30], the authors investigated a visualization technique, a tag cloud, to display tags to support search performance. They applied various dimensions to construct tag clouds for use in information retrieval usage. They explored parameters of constructing tag cloud layouts including font size, quadrant and proximity-to-largest-word during a presentation period or an interpretative period. The study showed that the list ordered by frequency is better for categorizing.

Another tag-based browsing construct is clustered tag clouds [41], which utilizes SOMs for visualization. The proposed approach not only facilitates the discovery of relationships between tags and corresponding content, but also improves tag-based navigation by clustering relevant tags. A similar idea, classified tag clouds, studied by Yahoo! Labs [33] classified tags by utilizing facets such as Wordnet. Their approach enabled Flickr photo browsing through different facets. Their analysis showed that users could effectively deploy query recommendations to explore large sets of images annotated with tags. Other studies [19, 35] explored another advanced tag construct, tag hierarchy, for tag-based navigation. By utilizing a decentralized search framework [35], the authors found that there are significant differences among different approaches to tag hierarchy construction in terms of success rate and average path length.

Since our primary goal intent in this paper is to explore whether the tag-based browsing constructs could provide any additional value to tag-based search, we apply the most popular interface layout, a tag cloud, as our basic tag interface and compare it to a traditional search box interface. Furthermore, according to our previous study [26] on image search, where we discovered that facets help users in exploring a large collection of images, we also investigate a faceted tag cloud interface in this study [33].

A similar study conducted by Sinclair and Cardew-Hall investigated the usefulness of tag clouds in terms of information seeking by analyzing the usage of tag clouds in a traditional search interface [34]. They found that subjects prefer tag clouds when the search task is more general, but favor issuing search queries, when more specific information is needed. Contrary to their study, our work is based on the domain of images where typically no descriptive content (such as page-text or abstract information) is given. Furthermore, we study three separate tag-based interfaces to discover the differences between a traditional search interface, a search interface enriched with tag clouds, and search interface extended with faceted tag clouds. In this setting, we can clearly identify *how people use* each interface and *how they perform*. To the best of our knowledge, this is the first work that compares multiple tag-based search interfaces.

4.8 Discussion and Conclusions

The main goal of the presented study was to perform a comparative user evaluation of tag-based browsing interfaces against simple search-based access to tagged collections. We compared user performance and feedback for three types of tag-based information access interfaces in the context of two recognized types of search tasks – lookup search and exploratory search. As expected, we obtained empirical evidence that the two tag-based browsing interfaces were superior to the baseline (search only) interface. At the same time, the analysis of objective data (performance and action profile) and of subjective data (questionnaires) produced slightly different results.

From the users' perspective, both tag-based browsing interfaces were perceived to be superior to the baseline. The users indicated that these interfaces provided significantly enhanced support for both types of user tasks and reported significantly higher levels of confidence that relevant information would be found. They also ranked both tag-based browsing interfaces significantly higher "overall" than the baseline interface.

From the performance and log analysis, significant differences were found for the traditional tag cloud interface when used in the exploratory search context. The tag cloud interface was found to be significantly more efficient in terms of both time and actions than the baseline interface. We also found that the tag cloud provided the most significant impact upon more difficult tasks and when the user was less familiar with the core topic of the task. A deeper analysis of user actions revealed another argument in favor of the tag cloud interface - with this interface, the "show more results" action was used significantly less often than in the baseline interface. This indicated that, with the tag cloud, the users were more likely to receive useful results at the top of the ranked list. None of these differences appeared to be significant for the faceted tag cloud; its objective performance was inferior to the performance of the traditional tag cloud. In addition, neither objective nor subjective data revealed any significant differences between the traditional tag cloud and the more advanced faceted tag cloud.

Why was the more advanced tag-based browsing interface less effective than the simpler tag-based browsing interface? Why was the faceted tag interface not a significant improvement over the baseline (search only) interface from a performance aspect? The post-session questionnaire provided some answers to these questions. This questionnaire asked users to select their "preferred" interface in light of two aspects : looking at performance in the past and looking forward to potential future uses of these interfaces. While the traditional tag cloud interface was preferred in previous tasks (which correlated with the objective performance data), the faceted tag cloud interface was the most popular for future use. It was also the top choice to be recommended to museum professionals. This was a strong indication that the faceted tag cloud interface was perceived as more powerful in the long run, but too difficult to use at first. This speculation is further confirmed by users' comments. In these comments, subjects stressed several aspects in which the faceted tag cloud interface was superior to the traditional cloud, yet indicated that it was harder to use at first. This data revealed that the faceted tag cloud interface should be assessed in a longer-term study, which would allow users to gain experience and become more proficient in operating with more sophisticated interfaces. We plan to explore this hypothesis in our future studies.

We also should acknowledge that the most noticeable differences observed in the study were not between the interfaces, but between the lookup and exploratory search tasks. Our data further confirmed that these two kinds of tasks are radically different. Exploratory search tasks are much harder; they consume more time and require more actions than lookup search tasks. Moreover, the very structure of user activities was very different between exploratory and lookup search. The occurrence of traditional search decreased considerably perhaps because it was much harder to find right keywords for the query. In contrast, almost 50% of user time in exploratory search context was spent on examining specific documents that were important to understand the domain and identify the most useful terms. These results correlate well with the previous research on exploratory search.

Finally, we should acknowledge a few limitations of our study. First, we focused on the query-to-image part of tag-based information access since it was the most different aspect among the three interfaces. The explicit presence of tags can also also enhance image-to-image navigation and further increase the value of tag-based browsing. Additional studies are required to determine the value of tags in this context. In addition, by the nature of our studies, we were unable to investigate one potential weakness of tag-based browsing in respect to classic search. All tag-based browsing interfaces require some considerable screen space for a tag cloud or other tag browsing artifact. This might reduce the space needed to show search results and decrease the effectiveness of tag-based browsing. In our studies, this effect was minimal: the study was performed on a regular desktop screen and search results were shown as thumbnails, which occupied relatively little space. We believe that, in this context, tag-based browsing interfaces were able to present a sufficient number of results despite the decreased presentation space. As the results shows, the Show More Results action was called upon significantly less frequently for the tag cloud. However, this might be of concern for those cases of mobile search with limited screen space as well as for different kinds of objects that require more space in the results presentation area. This is one of the reasons that we hesitate to generalize the observed results on tag-based information access to non-image resources. This is another aspect that requires additional investigation. We hope to explore some of these issues in our future work.

References

[1] P. , G. Koutrika, and H. Garcia-Molina. Can social bookmarking improve web search? *Proceedings of the international conference on Web search and web data mining*, pages 195–206, 2008.

[2] A. Ammari and V. Zharkova. Combining tag cloud learning with svm classification to achieve intelligent search for relevant blog articles. In *In Proceedings of the 1st International Workshop on Mining Social Media*, 2009.

[3] L. B. Baltussen. *Barbarians versus gatekeepers ? Tagging as a way of defining the ememrgent living archive paradigm*. PhD thesis, 2010.

[4] S. Bao, G. Xue, X. Wu, Y. Yu, B. Fei, and Z. Su. Optimizing web search using social annotations. In *Proceedings of the 16th international conference on World Wide Web - WWW '07*, page 501, New York, New York, USA, May 2007. ACM Press.

[5] S. Bateman, C. Gutwin, and M. Nacenta. Seeing things in the clouds: the effect of visual features on tag cloud selections. *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*, pages 193–202, 2008.

[6] K. Bischoff, C. S. Firan, C. Kadar, W. Nejdl, and R. Paiu. Automatically identifying tag types. In *Proceedings of the 5th International Conference on Advanced Data Mining and Applications*, ADMA '09, pages 31–42, Berlin, Heidelberg, 2009. Springer-Verlag.

[7] K. Bischoff, C. S. Firan, W. Nejdl, and R. Paiu. Can all tags be used for search? In *Proceedings of the 17th ACM conference on Information and knowledge management*, CIKM '08, pages 193–202, New York, NY, USA, 2008. ACM.

[8] D. Böhnstedt, L. Lehmann, C. Rensing, and R. Steinmetz. Automatic identification of tag types in a resource-based learning scenario. In *Proceedings of the 6th European conference on Technology enhanced learning: towards ubiquitous learning*, EC-TEL'11, pages 57–70, Berlin, Heidelberg, 2011. Springer-Verlag.

[9] K. Byström and K. Järvelin. Task complexity affects information seeking and use. *Inf. Process. Manage.*, 31:191–213, March 1995.

[10] K. Candan, L. Di Caro, and M. Sapino. Creating tag hierarchies for effective navigation in social media. *SSM '08: Proceeding of the 2008 ACM workshop on Search in social media*, pages 75–82, 2008.

[11] I. Cantador, I. Konstas, and J. M. Jose. Categorising social tags to improve folksonomy-based recommendations. *Web Semantics: Science, Services and Agents on the World Wide Web*, 9(1), 2011.

[12] E. H. Chi and T. Mytkowicz. Understanding the efficiency of social tagging systems using information theory. In *HT '08: Proceedings of*

the nineteenth ACM conference on Hypertext and hypermedia, pages 81–88, New York, NY, USA, 2008. ACM.

[13] A. Diriye, A. Blandford, and A. Tombros. When is system support effective? In *Proceedings of the third symposium on Information interaction in context*, IIiX '10, pages 55–64, New York, NY, USA, 2010. ACM.

[14] G. W. Furnas, C. Fake, L. von Ahn, J. Schachter, S. Golder, K. Fox, M. Davis, C. Marlow, and M. Naaman. *Why do tagging systems work?* ACM Press, New York, New York, USA, Apr. 2006.

[15] S. A. Golder and B. A. Huberman. The structure of collaborative tagging systems. *Journal of Information Science*, 32(2):198–208, 2006.

[16] Y. Hassan-Montero and V. Herrero-Solana. Improving tag-clouds as visual information retrieval interfaces. In *InScit2006: International Conference on Multidisciplinary Information Sciences and Technologies*, 2006.

[17] M. A. Hearst. *Search User Interfaces*. Cambridge University Press, 1 edition, 2009.

[18] M. A. Hearst and D. Rosner. Tag clouds: Data analysis tool or social signaller? In *HICSS '08: Proceedings of the Proceedings of the 41st Annual Hawaii International Conference on System Sciences*, Washington, DC, USA, 2008. IEEE Computer Society.

[19] D. Helic and M. Strohmaier. Building directories for social tagging systems. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, CIKM '11, pages 525–534, New York, NY, USA, 2011. ACM.

[20] D. Helic, C. Trattner, M. Strohmaier, and K. Andrews. On the navigability of social tagging systems. In *Proc. of 2010 IEEE International Conference on Social Computing*, pages 161–168, Los Alamitos, CA, USA, 2010. IEEE Computer Society.

[21] A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. Folkrank: A ranking algorithm for folksonomies. In *Proc. FGIR 2006*, 2006.

[22] Y. Kammerer, R. Nairn, P. Pirolli, and E. H. Chi. Signpost from the masses: learning effects in an exploratory social tag search browser. In *Proceedings of the 27th international conference on Human factors in computing systems*, CHI '09, pages 625–634, New York, NY, USA, 2009. ACM.

[23] O. Kaser and D. Lemire. Tag-Cloud Drawing: Algorithms for Cloud Visualization. *Proceedings of Tagging and Metadata for Social Information Organization (WWW 2007)*, 2007.

[24] D. Kelly. Methods for evaluating interactive information retrieval systems with users. *Found. Trends Inf. Retr.*, 3:1–224, January 2009.

[25] K. Knautz, S. Soubusta, and W. G. Stock. Tag clusters as information retrieval interfaces. In *Proceedings of the 2010 43rd Hawaii International Conference on System Sciences*, HICSS '10, pages 1–10, Washington, DC, USA, 2010. IEEE Computer Society.

[26] Y. Lin, J.-W. Ahn, P. Brusilovsky, D. He, and W. Real. Imagesieve: Exploratory search of museum archives with named entity-based faceted browsing. *Proceedings of the American Society for Information Science and Technology*, 47(1):1–10, Nov. 2010.

[27] G. Marchionini. Exploratory search: from finding to understanding. *Commun. ACM*, 49:41–46, April 2006.

[28] J. Oomen, L. B. Baltussen, Sander Limonard, M. Brinkerink, A. van Ees, L. Aroyo, J. Vervaart, K. Afsar, and Riste Gligoro. Emerging Practices in the Cultural Heritage Domain - Social Tagging of Audiovisual Heritage. In *Web Science10:Extending the Frontiers of Society On-Line*, Raleigh, NC, USA., 2010.

[29] S. Overell, B. Sigurbjörnsson, and R. van Zwol. Classifying tags using open content resources. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, WSDM '09, pages 64–73, New York, NY, USA, 2009. ACM.

[30] A. Rivadeneira, D. Gruen, M. Muller, and D. Millen. Getting our head in the clouds: toward evaluation studies of tagclouds. *Proceedings of the SIGCHI conference on Human factors in computing systems*, page 998, 2007.

[31] A. W. Rivadeneira, D. M. Gruen, M. J. Muller, and D. R. Millen. Getting our head in the clouds: toward evaluation studies of tagclouds. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 995–998, New York, NY, USA, 2007. ACM.

[32] S. Sen, S. K. Lam, A. M. Rashid, D. Cosley, D. Frankowski, J. Osterhouse, F. M. Harper, and J. Riedl. tagging, communities, vocabulary, evolution. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work CSCW 06*, volume 4 of *CSCW '06*, pages 181–190. ACM, ACM Press, 2006.

[33] B. Sigurbjörnsson and R. van Zwol. Tagexplorer: Faceted browsing of flickr photos. In *Technical Report, Yahoo! Research*, 2010.

[34] J. Sinclair and M. Cardew-Hall. The folksonomy tag cloud: when is it useful? *Journal of Information Science*, 34:15, 2008.

[35] M. Strohmaier, D. Helic, D. Benz, C. Körner, and R. Kern. Evaluation of folksonomy induction algorithms. *ACM Trans. Intell. Syst. Technol.*, 2012.

[36] C. Trattner, C. Körner, and D. Helic. Enhancing the navigability of social tagging systems with tag taxonomies. In *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, i-KNOW '11, pages 18:1–18:8, New York, NY, USA, 2011. ACM.

[37] C. Trattner, Y.-L. Lin, D. Parra, W. Real, and P. Brusilovsky. Evaluating tag-based information access in image collections. In *Proceedings of the 23rd ACM Conference on Hypertext and Social Media*, HT '12, New York, NY, USA, 2012. ACM.

[38] P. Vakkari, M. Pennanen, and S. Serola. Changes of search terms and tactics while writing a research proposal a longitudinal case study. *Inf. Process. Manage.*, 39:445–463, May 2003.

[39] P. Venetis, G. Koutrika, and H. Garcia-Molina. On the selection of tags for tag clouds. In *Proceedings of the fourth ACM international conference on Web search and data mining*, WSDM '11, pages 835–844, New York, NY, USA, 2011. ACM.

[40] C. Wartena. Automatic classification of social tags. In *Proceedings of the 14th European conference on Research and advanced technology for digital libraries*, ECDL'10, pages 176–183, Berlin, Heidelberg, 2010. Springer-Verlag.

[41] A. Zubiaga. Enhancing Navigation on Wikipedia with Social Tags. *Wikimania 2009*, pages 1–9, 2009.

Part III

Evaluation II: Comparing Tags with Tag-Alike Meta-Data

To what extent are tags/tag clouds more useful/efficient for search/navigation than other tag-alike meta-data such as keywords or search query-terms?

5 Linking Related Content in Web Encyclopedias with Search Query Tag Clouds

This chapter is based on the paper *"Linking Related Content in Web Encyclopedias with search query tag clouds"* published in the IADIS International Journal on WWW/ in 2011.

Another interesting issue in the context of tag-based search and navigation is the question as to what extent tags are more useful for navigation than tag-alike meta-data structures. Since tags are very related to the notation of keywords and since related research has shown that tags are in their structure comparable to so-called query tags harvested from search query logs [5], we are interested in studying the navigational efficiency of tags compared to these tag-alike meta-data structures. To that end, we introduce in this chapter a paper that presents a tool called QueryCloud which enables an information system to link related content over tag clouds that are created from historic search queries. By comparing the approach on a theoretical and empirical level with a tagging system maintained by real users of an information system called the Austria-Forum we show that tag clouds based on query terms out-perform tag clouds created from tags collected from real users in terms of linking and efficiently navigating related content.

The original contribution can be found in [25].

5.1 Abstract

In this paper we present a novel tool for exploring related resources in Web encyclopedias called QueryCloud. Typically, users come to an encyclopedia from a search engine such as Google, Yahoo! or Bing and upon reading the first page on the site they leave it immediately thereafter. To tackle

this problem in other systems such as Web shops, additional browsing tools for easy finding of related content are provided. In order to overcome this issue in the context of Web encyclopedia systems, we introduce a tool called QueryCloud. The tool combines two promising approaches - tag clouds and historic search queries - into a new single one, i.e. each document in the system is enriched with a tag cloud containing collections of related concepts populated from historic search queries. To test the feasibility of the approach, we integrated a prototypical implementation of the tool into a large Web encyclopedia called the Austria-Forum and conducted several experiments on a theoretical and empirical level. As our experiments show, QueryCloud provides a great alternative to traditional forms of tag cloud creation. With several experiments on a theoretical and empirical level we show that tag clouds generated by our system out-perform tag clouds that are based on user-tags in terms of linking content and navigability.

5.2 Introduction

Nowadays, content in Web encyclopedias such as Wikipedia is mainly accessed through search engines (Wikimedia 2010). Typically, users with a certain interest in mind go to a search engine such as Google, Yahoo! or Bing, define a search query there and click on a link from the result list from which they are referred to an article within Wikipedia. Upon reading the document they decide to either go back to the search engine to refine their search, or close their browser if they have already found the information they needed. Such a user behavior on encyclopedia sites is traceable through a typical high bounce rate (Alexa 2010, [7]. Essentially, users do not "really" browse in online encyclopedia systems such as Wikipedia to find further relevant documents [7] - they rather use search engines such as Google, Yahoo! or Bing for that purpose. It is our opinion that Web encyclopedias simply lack usable tools that support users in explorative browsing or searching. For example, in Web based systems such as Web shops, different approaches have been applied to tackle this situation. Amazon for instance offers the user related information through collaborative filtering techniques for each product. Google or Yahoo! apply a similar approach by offering related content (e.g. sponsored links) to the user by taking the users' search query history into account [18]. On the other hand, social tagging has emerged as an interesting alternative to find relevant content on the Web [11]. These systems apply the concept of social navigation [19], i.e. users browse by means of tag clouds, which are collections of keywords assigned to different online resources by different users [11] driven by different motivations [21, 2].

In this paper we introduce a novel tool called QueryCloud to offer related content to users of Web encyclopedias. Essentially, the tool is based on the simple idea of integrating a tagging system into a Web encyclopedia and offering related content to users via the so-called resource-specific tag clouds by automatically linking related documents over the users search query terms. In this way two promising approaches are successfully combined into a new single one - tag clouds and historic search queries. To test the approach, we implemented a prototype of the tool and integrated it in a large Web encyclopedia called the Austria-Forum. To evaluate the system by means of link quality, tag network quality and navigability, we conducted several experiments on a theoretical level. Additionally to this, we conducted a user study to investigate whether or not the tags generated by our system are also meaningful for the user by describing the content of a particular Web page. Hence, the overall contribution of this paper can be summarized as follows:

- Introduction of a novel tool called QueryCloud that links related content in Web encyclopedia system via so-called resource specific search query tag clouds.
- Evaluation of the tool by integrating it into a large Web encyclopedia system called the Austria-Forum and comparing it on a theoretical and empirical level with tags and tag clouds that are based on tags generated by real users of this system.

Essentially, the paper is structured as follows: In Section 5.3 we present the basic idea of this new approach. Section 5.4 shortly discusses the implementation of the novel tool. In Section 5.5 we provide an analysis of the potentials and limitations of our tool. Section 5.6 discusses related work. The final Section 5.7 concludes the paper and provides insights in the current progress of the project.

5.3 Approach

The basic idea of QueryCloud is to offer related content to users via the so-called resource-specific tag cloud by automatically linking related documents via the search query terms of the users. On the one hand, tag clouds represent an interesting alternative navigation tool in modern Web-based systems. Moreover, they are very close to the idea of explorative browsing [22], i.e. they capture nicely the intent of users coming to a system from a search engine - users have searched in e.g. Google, Yahoo! or Bing and now they click on a concept in a tag cloud that reflects their original search intent. On the other hand, search query history, i.e. queries that are “referrers” to found documents are an invaluable source of information for refining user search in the system. It is our belief that an integration of such a tool online encyclopedia systems would greatly contribute to leading users to related documents. In order to make the idea

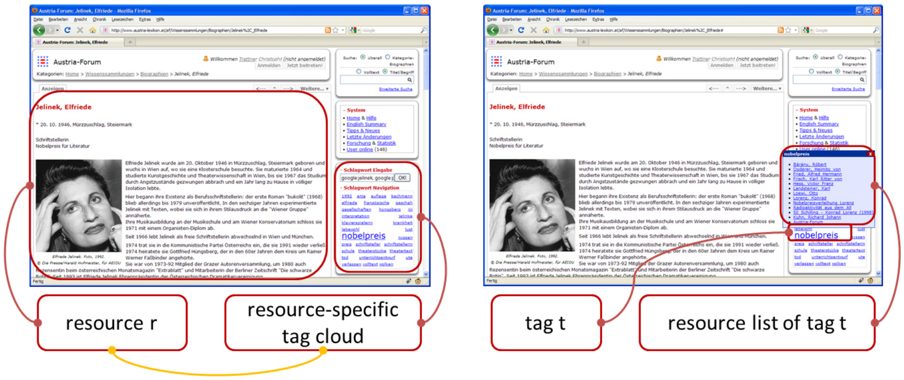


Figure 5.1: Example of resource-specific query tag cloud and resource list of a tag within Austria-Forum rendered by QueryCloud.

work the users search query history needs to be obtained in a resource specific way. This is achieved by collecting the HTTP-Referrer information from a user coming from a search engine to a particular website (resource) within a Web based encyclopedia system such as Austria-Forum. To retrieve the user's HTTP-Referrer information, we use in QueryCloud the simple approach of a JavaScript code snippet (see Tag Collection Module in Section 3) that has to be included by the owner of the particular Web site. Now, how does the script work? Let us give an example: Suppose a user goes to Google and searches for "Elfride Jellinek Biographie" and selects the link <http://www.austria-lexikon.at/af/Wissenssammlungen/Biographien/Jelinek%2C_Elfriede> from the results list that refers to the Elfriede Jelinek biography in the Austria-Forum. The QueryCloud Tag Collection Module then simply parses the HTTP-Referrer <http://www.google.at/search?hl=de&q=elfriede+jellinek+biographien> information of this site and extracts from it the user's search query terms "elfride", "jellinek", "biographien". The tags are then stored in the QueryCloud Tag Store Module (see Section 5.4). This procedure is performed whenever a user lands on a Web page from a search engine within the Austria-Forum. The tags are then used to create a resource-specific tag cloud for each site. This is done by the QueryCloud Tag Cloud Generation Module (see Section 5.4). The Tag Cloud Presentation Module (see Section 5.4) renders the tag cloud then in a visually appealing fashion and presents it to the user. Hence, two pages (or even more) are linked with each other by this approach if they have the same query tag in common. Upon clicking on a particular query tag in a resource-specific tag cloud the user is provided with a list of links of the resources which have this query tag in common (see Figure 5.1). By clicking on a particular link in the resource list the user is then forwarded to the resource she was searching for (see Figure 5.2).

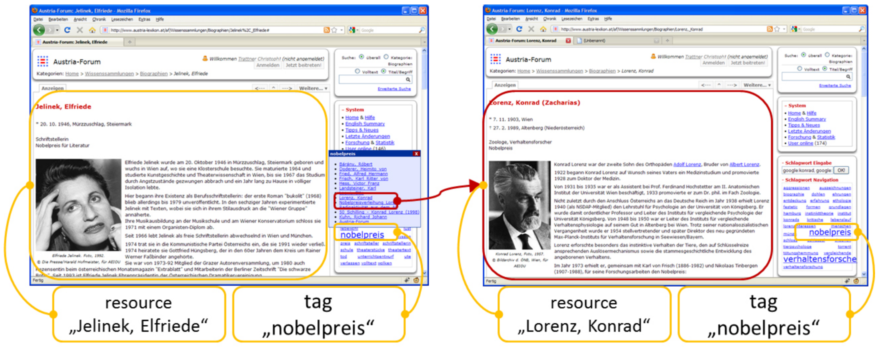


Figure 5.2: Example of tag cloud driven navigation within Austria-Forum using QueryCloud.

5.4 Implementation

The first prototypical implementation [27] consists of four independent modules (see Figure 5.3).

Tag Collection Module: The tag collection module is the first module within the QueryCloud system. Basically, this module is a simple client part module which retrieves HTTP-Referrer information, time information and target page of a user coming from a search engine such as Google, Yahoo! or Bing to a website.

Tag Storage Module: This module is the actual heart peace of the system. It provides a couple of interface routines for storing and deleting data from the database back-end module that is implemented with Apache Lucene. The module also integrates a number of clean-up and stemming routines for eliminating stop-words, punctuation characters or plural words. The synonym problem is not considered by this module. Due to reasons of performance two index files are generated: One for expressing the tag resource-relations $(t_n \| r_i, ..., r_j)$ and one for expressing the resource-tag relations $(r_n \| t_i; ..., t_j)$. In this way, tags and resources can be searched independently from each other, to either create a tag cloud for a specific resource or to create a resource list for a specific tag.

Tag (Cloud) Generation Module: To provide the access to related documents a resource-specific search query term/tag cloud is calculated by this module. This tag cloud is of the form $TC_r = (t_1, ..., t_n, r_1, ..., r_m)$, where $r_1, ..., r_m$ are the resources which have any of the query tags $t_1, ..., t_n$ in common. For retrieving the query tags and the corresponding resources (cf. Figure 1), this module provides a simple HTTP interface using the following

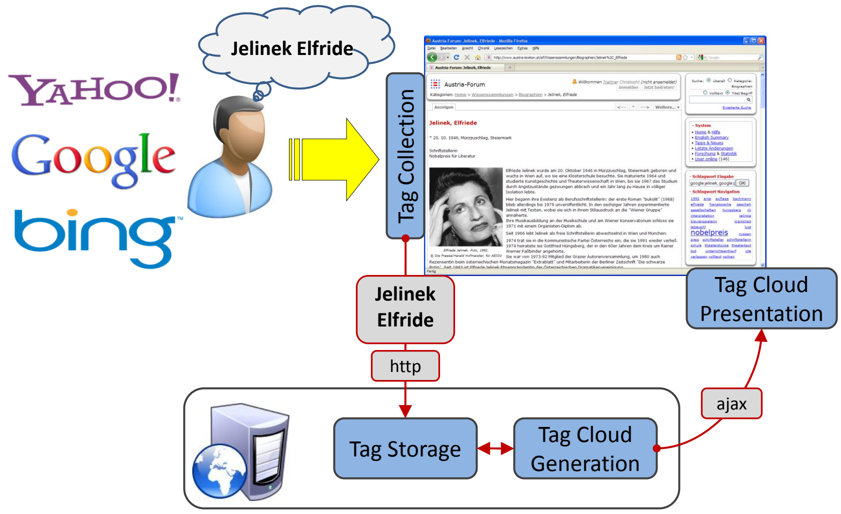


Figure 5.3: QueryCloud - structural diagram.

two functions:

- GetTagCloud(URL,max. tag cloud size) generates a XML representation of a query tag cloud
- GetResources(URL, tag, max. resource list length) generates a XML representation of the resource list for a particular query tag

Tag Cloud Presentation Module: This module is a client-side AJAX module implemented in JavaScript. It retrieves the XML representation of a query term/tag cloud or an XML representation of a resource list of a particular query term/tag from the tag cloud generation module and renders a tag cloud in a visually appealing fashion.

5.5 Evaluation

To investigate the feasibility of the tool before actually deploying it, we first decided to only integrate the QueryCloud tag collection module into the Austria-Forum life system to collect the search query terms from users coming from search engines such as Google, Bing or Yahoo! to the Austria-Forum for a period of 240 days. On the basis of this dataset, we conducted several experiments on a theoretical and empirical level to evaluate the feasibility of the approach. In the following subsections we will give a short description of our test system, the tag data sets used in the experiments, a detailed description of the experiments and their results.

	#res	#tags	#tas
AF Dataset	13,398	11,516	33,737

Table 5.1: AF dataset: Number of resources (#res), number of tags (#tags) and number of tag assignments (#tas).

5.5.1 Test System

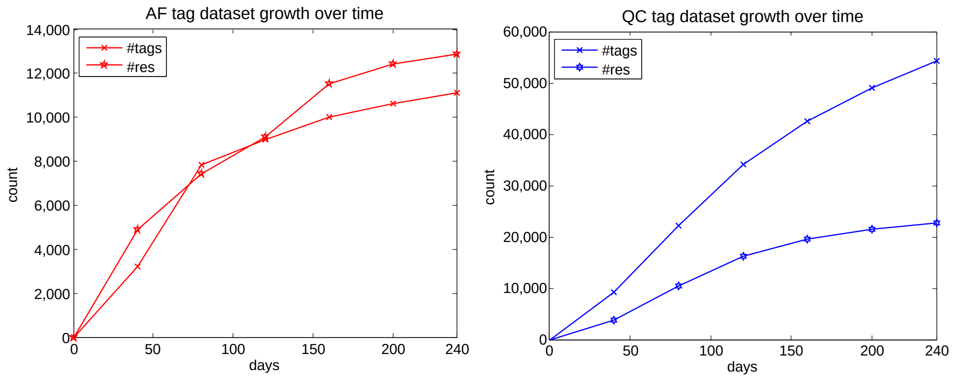
As described before, we used a Web encyclopedia called the Austria-Forum [26] as our test system to investigate the feasibility of our approach. Basically, Austria-Forum is a wiki-based online encyclopedia containing articles related to Austria. The system comprises a very large repository of articles, where new articles are easily published, edited, checked, assessed, and certified, and where the correctness and quality of each of these articles is assured by a person that is accepted as an expert in a particular field. As of July 2010, the system contains more than 130,000 information items, attracts more than 4,000 distinctive users (over 80% coming from search engines such as Google, Bing or Yahoo!) each day and is known as the biggest online encyclopedia system on the Web containing content about Austria [26].

5.5.2 Baseline

As baseline for our experiments a tag dataset (further referred as the AF dataset) of the Austria-Forum was used. Since the tags are generated by real users of the system to describe or categorize [15] the contents, we used this dataset as the baseline to compare our QueryCloud approach with. As of February 2010, the AF tag dataset contains 11,516 tags (#tags), 13,398 resources (#res) and 33,737 tag assignments (#tas) (see Table 5.1).

5.5.3 Measuring Tag Quantity

Since the success of the whole concept depends on automatically applying tags to the resources of a Web based encyclopedia system we conducted in the first step an experiment that measured tag quantity. In particular, we investigated the growth of the number of resources (#r), the growth of the number of generated tags (#t) and the number of tag assignments (#tas) in general for a period of 240 days (cf. [28]). Essentially, we could observe that QueryCloud generated on average 226 new tags (#t) per day and annotated on average 95 new resources every day (#r) (see Figure 5.4). The average number of tagged resources was 1290 per day. Compared to this, the Austria-Forum taggers annotated 135 resources on average per day, generated on average 46 new tags (#t) and annotated on average 53 new resources (#r) every day, for the time the tagging functionality was first introduced in the Austria-Forum (see Figure 4). Hence, the QueryCloud



(a) AF **(b)** QC

Figure 5.4: Number of tagged resources (#res) and number of generated tags (#tags) over time for the Austria-Forum: QC dataset (Figure a) on vs. AF dataset (Figure b).

tagging approach clearly outperforms the Austria-Forum human taggers by annotating 9 times more resources, and generating 5 times more new tags and annotating 1.7 times more resources in the same period of time. As of July 2010, the QC tag dataset (further referred as QC dataset) contains 54,379 tags (#tags), 22,798 resources (#res) and 309,683 tag assignments (#tas) (see Table 5.2).

5.5.4 Measuring Link Quality

After measuring the quantity of the tags produced by QueryCloud system we had a closer look at the actual “link quality” of the produced tags by QueryCloud system. Since the success of the whole concepts depends on linking related documents over tags that share more than one resource with each other, we conducted an experiment measuring the number orphan tags produced by QueryCloud system. Orphan tags are basically tags which are applied to only one resource within a tagging system, i.e. they do not connect any resources with each other. Again, we could observe that QueryCloud system performs really well by actually producing 7% less orphan tags than the human taggers (AF) do (see Table 5.3). Note, for the rest of our experiments we only used the cleaned datasets AF and QC, cleaned from orphan tags and their corresponding resources. In Table 5.4 the new statistics of both datasets are presented.

	#res	#tags	#tas
QC Dataset	22,798	54,379	309,683

Table 5.2: QC dataset: Number of resources (#res), number of tags (#tags) and number of tag assignments (#tas).

	#tags	#orphans
QC Dataset	54,379	34,962 (70%)
AF Dataset	11,516	8,865 (77%)

Table 5.3: Number of tags (#tags) and number of orphan tags (#orphans): QC dataset vs. AF dataset.

5.5.5 Measuring Network Quality

Another metric we were interested in was the so-called "network quality" of the QueryCloud system. In order to measure this property, we first modeled the QueryCloud tag cloud network as a simple tag-resource bipartite graph system of the form $V = R \cup T$, where R is the resource set and T is the query tag set [9]. Hence, for the following experiments, we assume that neither the tag cloud size nor the resource list is limited, which, as shown in one of our own previous works influences tag cloud navigability [10]. For more details, see [24] for a generic solution of that problem, to create efficiently navigable tag cloud networks with limited resource lists. Since the "link quality" experiment only showed us how many actual useful tags the systems generates (by means of connecting two or more resources with each other), we investigated in this experiment how many resources are connected with each other. To measure this metric, we calculated the connected components [17]. Essentially, we could observe that size of the largest connected component in the QC dataset is 99%, i.e. 99% of all resources within Austria-Forum are connected via tag clouds generated by the QueryCloud system (cf. [28]). Contrary to this, the AF dataset generates a tag cloud network which is only to 94 % connected (cf. [28]).

5.5.6 Measuring Navigability

In the fourth experiment, we examined the property of the tool to navigate to related documents within a Web based encyclopedia system. In [10] we have shown that navigable tag cloud networks have certain properties. According to Kleinberg [13, 12, 14] a navigable network can be formally defined as network with a low diameter [20] bounded poly-logarithmically, i.e. by a polynomial in log(N), where N is the number of nodes in the network, and an existing giant component. Thus, as a first step, we examined again the size of the largest connected component. As shown in the previous section, QueryCloud generates a tag cloud network whose largest connected component contains

	#res	#tags
QC Dataset	18,831	11,485
AF Dataset	12,103	2,207

Table 5.4: Orphan cleaned up dataset statistics of QC dataset and AF dataset.

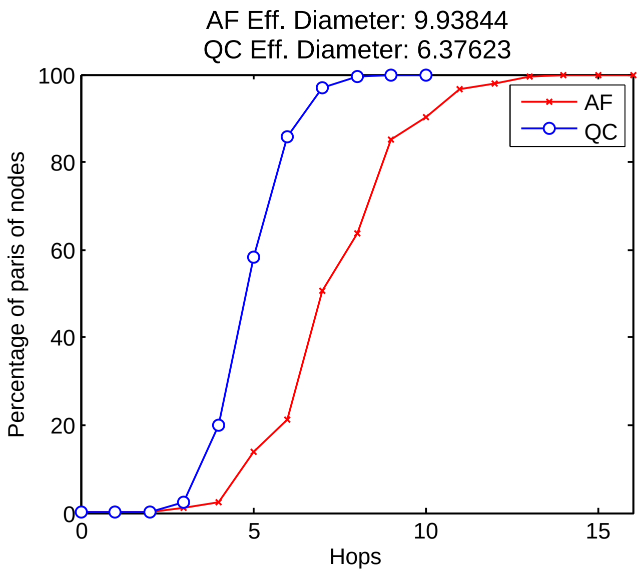


Figure 5.5: Distribution of shortest path pair lengths for QC tag cloud network (blue line) and AF tag cloud network (red line).

almost all nodes (99%) of the network. Contrary to this, the AF dataset generates a tag cloud network which is "only" connected to 94%. Thereafter, we calculated the number of shortest path pairs within QueryCloud's tag cloud network and investigated the effective diameter of the network. As Figure 5.5 shows, QueryCloud generates a tag cloud network whose effective diameter is around 6.3 hops while the AF dataset generates a tag cloud network with an effective diameter of around 9.9 hops. Putting the results of these two experiments together we can see that QueryCloud produces a navigable tag cloud network, $log_2(18,831) = 14.2 > 6.3$ [9, 13, 12, 14]. The same applies to the AF tag cloud network, since $log_2(12,103) = 13.6 > 9.9$.

5.5.7 Measuring Efficiency

Now, since we have shown that tag cloud network of QueryCloud system is navigable, we also wanted to know how "efficiently" navigable the tag cloud network is in general. As shown by Kleinberg, an "efficiently" navigable network is a network possessing certain structural properties so that it is

Algorithm 2 Hierarchical Decentralized Searcher [8]

INPUT: Tag-Resource Graph G, Start node $START$, Target node $TARGET$
/* returns a tag-taxonomy based on degree centrality and tag co-occurance [8]*/

$T \leftarrow GetDegCoocTagTaxonomy(G)$
$currentNode \leftarrow START$
while $currentNode! = TARGET$ **do**
  $neighbors \leftarrow getalladjacentnodes \in Gfromcurrent Node$
  /*finds closest node according to dist = min, where dist(A,B) = h(A)+h(B)-2h(A,B)-1*/
  $currentNode \leftarrow findClosestNode(neighbors, T)$
end while

possible to design an decentralized search algorithm, i.e. an algorithm that only has local knowledge of the network, and whose delivery time (the expected number of steps to reach an arbitrary target node) is poly-logarithmic or at most sub-linear in N, where N is the number of nodes in the network [13, 12, 14]. Essentially, we implemented a decentralized searcher (see Algorithm 2) based on the ideas of [1] to evaluate the actual efficiency of the system to navigate to related documents using tag clouds for navigation [8]. In order to support efficient search, we utilized a hierarchical background knowledge base (in our case a tag-taxonomy) to find related content within a tag cloud network (see Algorithm 2). As appropriate tag taxonomy, we us a tag-taxonomy which is based on the so-far best known tag-taxonomy induction algorithm (creating best semantic relations and search results) called Deg/Cooc [23].

To find a certain resource A (e.g. tagged as "spätromantik") from a certain node (e.g. tagged as "schloß") within the network (see Figure 5.6 and Figure 5.7), the searcher first selects all adjacent nodes for the start node and then selects the node from the network ("kirche") which has the shortest distance $dist(A, B) = h(A)+h(B)-2h(A, B)-1$ to target node B in the tag-taxonomy, with $h(A)$, $h(B)$ being the heights of the two nodes A, B in the hierarchy and with $h(A, B)$ being the height of the least common ancestor of the two nodes A, B in the hierarchy [1, 8]. This process is continued until node B is reached. In order to get statistically significant results, we simulated 100,000 user-search-requests randomly requesting for a resource B within the system starting at randomly selected resource A. As shown in Figure 5.8, we can observe that with the help of QueryCloud system a user is able to find a resource within Austria-Forum in an efficient way, i.e. within only 8 hops almost 98% of all resources in the Austria-Forum can be reached, $log_2(18, 831) = 9.8 > 8$. Contrary to this, the AF tag dataset provides successful finding of related resources in the Austria-Forum only in 68% of the cases and this also in significantly more steps (14 hops) than with the QC dataset.

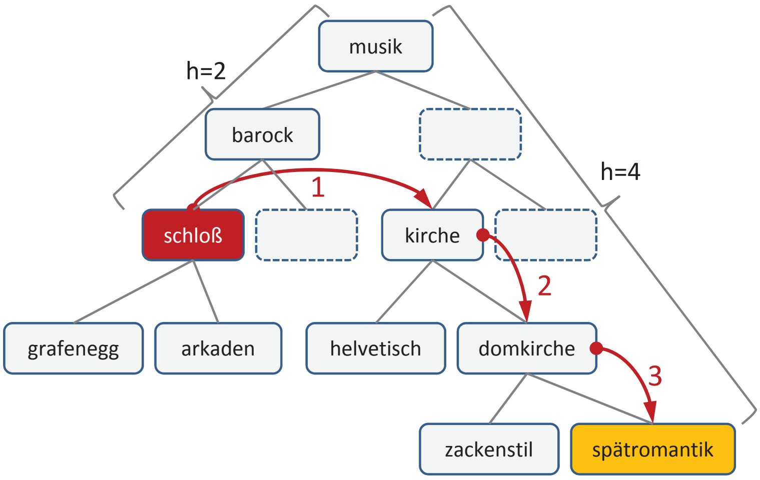


Figure 5.6: Example of a tag-taxonomy generated for AF dataset.

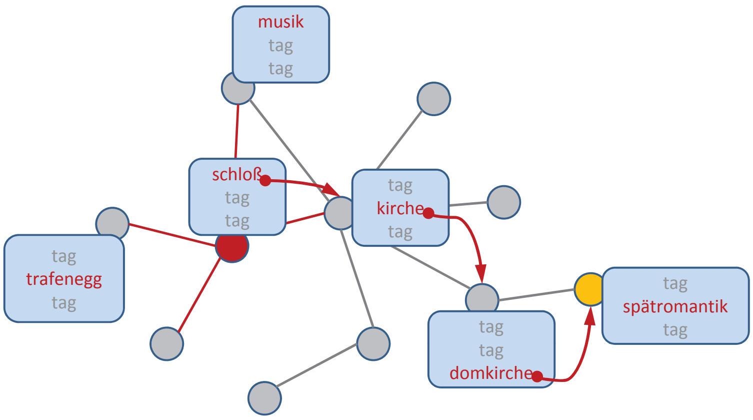


Figure 5.7: Example of a resource-specific tag cloud network and a search through it utilizing the tag-hierarchy from Figure 5.6 as background knowledge.

5.5.8 Measuring Tag Quality

Last but not least, we investigated in our final experiment the quality of the tags, respectively the tag clouds generated by our QueryCloud system. Even if related work has shown that query tags can be utilized to generate shorter navigational paths between documents [4] or that they are very similar to tags generated by the users [6], none the previous studies has shown whether or not query tags also provide a meaningful source for the users to describe

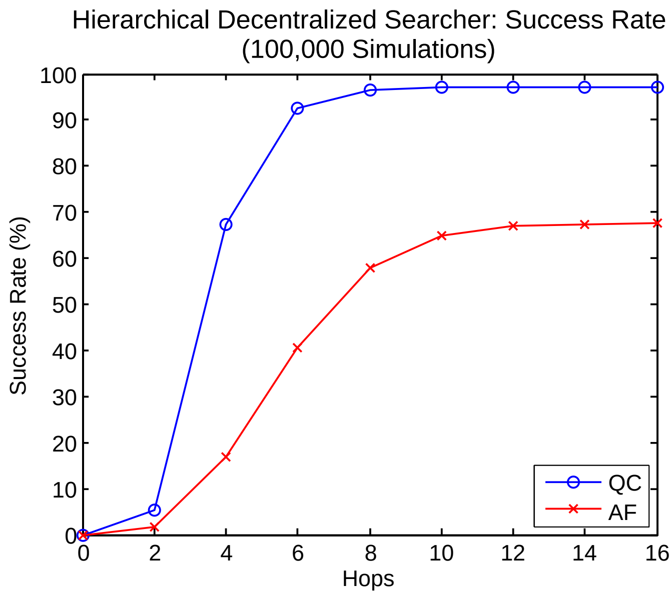


Figure 5.8: Success Rate plot for the hierarchical decentralized searcher in the QC tag cloud network (blue line) and AF tag cloud network (red line).

Web content. For that reason, we conducted a user study where we asked our test subjects to find out of a set of query tags the tags which were not relevant for a given Web page in the Austria-Forum system. As baseline for the experiment we again used the user generated tags from the AF tag dataset. The reason why we did not ask our users to explicitly find the relevant terms out of the set of query tags was basically the fact that we observed in a pilot study that users had more problems in highlighting items which are relevant than finding items which are *not relevant* for them for a given resource. Furthermore, was the number of not relevant items small than the number of relevant items which made it easier for the user to just point out the non-relevant items and to complete the online questionnaire in a faster. All in all, the setup of the user study was the following:

1. In the first step, we selected uniform at random, 250 resources overlapping resources from both datasets (QC and AF). In order to get meaningful results, only those resources were taken into account for which at least two tags (query tags and user tags) were present per resource. For the samples we took, the size of the tag cloud for the QC dataset was 5.2 terms on average and 2.7 terms on average for the AF dataset.

2. After that, we combined the tags of both datasets. We did that since we wanted our users not to know whether they are high lightening a

	NRR	p
QC Dataset	4.8%	0.376
AF Dataset	2.4%	

Table 5.5: Mean non-relevance feedback rating (=NRR) for QC dataset and AF dataset over all test users and p-value for the two sample t-test.

query tag as relevant for a given resource in Austria-Forum or a real user tag. The overlap between the QC dataset and the AF dataset was in general not very high. Contrary to the findings of [6] we can find only 9% of the query tags in the AF dataset and 16% of the user generated tags in the QC dataset.

3. In the third step, we implemented an online study with 50 tasks (=resources) per questionnaire. Overall we created five different questionnaires which covered the sample of 250 resources we have chosen from the Austria-Forum system.

4. Finally, we set up the user study on one of our servers in the TU-Graz domain and invited colleagues and friends to participate in the study. The whole study started at December 1st 2011 and was online for one week.

Overall, we had 15 test subjects from three different departments of our university who participated in the experiment. The test subjects were randomly assigned to each questionnaire. The mean length of time to complete a questionnaire was 35 minutes. For the final evaluation, we took only those tags into account where at least all three test users had an agreement on. In order to conduct the overall rating quality of our test subjects, we calculated the inter-rater agreement (=k) of the users according to Fleiss'kappa. The mean inter-rater agreement of the users for the query tags was k=0.22 and for the user tags it was k=0.21. According to the inter-rater agreement levels of Landis and Koch [16] this can be interpreted as fair agreement (k=0.21-0.40).

In Table 5.5 the mean non-relevance feedback ratings (=NRR) for the two datasets QC and AF of the user test are presented. As shown, we can observe overall a very small number of tags in both datasets which were rated as non-relevant. For the QC dataset sample the mean non-relevance rating was 4.8% and for the AF dataset sample it was 2.4%. To see whether or not the two values were also statistically significantly different, we performed two-sample t-test. As also shown in Table 5.5, on a confidence interval of 95%, the differences were not significant, $p = 0.376$. Even if the experiment did not show that query tags generated by QueryCloud out-perform user generated tags by means of term relevance, the study revealed that (even if query tags are not identical to user tags) they are to a high degree relevant for the user. Overall, we could show that query tags generated by QueryCloud are almost to the same degree relevant for the user of a given Web page as user generated tags.

5.6 Related Work

The almost first notable work in the field of search queries and user tags was a study conducted by Antonellis et al. [3]. In their work the authors performed a set of experiments to study the information value of search engine queries when treated as "tags" or "labels" for the Stanford domain. In particular, the authors tried to answer the question how much extra information these query tags provide for web pages by comparing them to tags from the Delicious bookmarking site and to the page text. As datasets for their experiments the authors used a self-collected query log dataset retrieved over the users HTTP-Referrer Information and a crawled Delicious tag dataset for the stanford.edu domain. The authors conclude their work, that query tags can provide substantially many (on average 250 tags per URL), new tags (on average 125 tags per URL are not present in the page text) for a large fraction of the stanford.edu domain.

In another study which was conducted in the same year, Carman et al. [6] investigated tags and query logs to see whether or not the terms people use to annotate websites are also similar to the websites search query terms. Interestingly, the authors found out that the vocabulary used for tagging and search is quite similar, however not identical. Additionally to these findings, the authors tried to answer the question whether or not search queries are more related to page content than tags generated by users. In a number of experiments, they found out that query tags are more related to page content than user tags. As datasets for their experiments the authors used the AOL search query log and tags from Delicious.

Another relevant work is the paper "Query Logs as Folksonomies" conducted by Benz et al. [5] in 2010. In their study the authors investigate the extent to which folksonomies can be generated from query log files. The focus is on three comparative studies of the system's content, structure and semantics. The results show that query logs incorporate typical folksonomy properties and that approaches to leverage the inherent semantics of folksonomies can be applied to query logs as well.

In a very interesting follow-up work [4] of Antonellis et al., the authors describe how they built a navigational tool on the basis of query tags. Additionally to this, they proposed a framework for comparing different tag selection methods. Similar to our own work they investigated user tags and query tags by means of the so-called navigational utility. Put simply, they calculated the navigational utility of a page as the number of resources that could be reached through this page and the closeness to other resources based on the pages tags. The most interesting finding in their work was that based on the measure of the navigational utility of a page, query tags increase the navigational utility of a page more than tags extracted from tags or user tags, or in other words, query tags are a better source for navigational tags than tags extracted from text or tags assigned by users. Their findings actually go along with the results as we also observed them in this paper. However, contrary to our own work, the authors focused in their work only on the tag/link selection problem to increase the navigational utility of a page rather than focusing on the general question to what extent query tag clouds outperform user generated tag clouds by means of tag quantity, link/network quality, navigational efficiency and quality, as we did in this paper.

The last work to be mentioned in this field is a study that was actually conducted by us in 2010. In particular, in [28] we presented the idea of combining search query tags and user tags to increase the navigability of tagging systems. To that end, we introduced measures such as tag cloud coverage and the size of the largest strongly connected component. Overall, we could show that the navigability of tagging systems could be significantly increased if we enrich an existing tagging system with query log terms.

5.7 Conclusions

In this paper we presented a novel tool called QueryCloud for exploring related resources in Web encyclopedias. The tool aims at to offering additional navigational paths to related resources for users of such systems in general, and for users who come to these systems from search engines such as Google, Yahoo! or Bing in particular. Furthermore, we showed the potentials and limitations of the tool by integrating it into a large Web encyclopedia system called the Austria-Forum. By comparing QueryCloud on a theoretical and empirical level with tag clouds that are based on tags generated by real users of the Austria-Forum we showed that our system out-performs these tag clouds in terms of linking and navigating related content. Additionally to this, we could show in a user study that query tags generated by QueryCloud are almost to the same degree relevant for the user of a given Web page as user generated tags.

References

[1] L. Adamic and E. Adar. How to search a social network. *Social Networks*, 27(3):187 – 203, 2005.

[2] M. Ames and M. Naaman. Why we tag: Motivations for annotation in mobile and online media. In *Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI 2007)*, CHI'07, pages 971–980, USA, NY, 2007. ACM.

[3] I. Antonellis, H. Garcia-Molina, and J. Karim. Tagging with queries: How and why? In *WSDM (Late Breaking-Results)*, 2009.

[4] I. Antonellis, J. Karim, and H. Garcia-Molina. Navigating the web with query tags. Technical report, Stanford University, 2010.

[5] D. Benz, A. Hotho, R. Jäschke, B. Krause, and G. Stumme. Query logs as folksonomies. *Datenbank-Spektrum*, 10(1):15–24, 2010.

[6] M. J. Carman, M. Baillie, R. Gwadera, and F. Crestani. A statistical comparison of tag and query logs. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '09, pages 123–130, New York, NY, USA, 2009. ACM.

[7] D. F. Gleich, P. G. Constantine, A. D. Flaxman, and A. Gunawardana. Tracking the random surfer: empirically measured teleportation parameters in pagerank. In *Proceedings of the 19th international conference on World wide web*, WWW '10, pages 381–390, New York, NY, USA, 2010. ACM.

[8] D. Helic and M. Strohmaier. Building directories for social tagging systems. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, CIKM '11, pages 525–534, New York, NY, USA, 2011. ACM.

[9] D. Helic, C. Trattner, M. Strohmaier, and K. Andrews. On the navigability of social tagging systems. In *Proc. of 2010 IEEE International Conference on Social Computing*, pages 161–168, Los Alamitos, CA, USA, 2010. IEEE Computer Society.

[10] D. Helic, C. Trattner, M. Strohmaier, and K. Andrews. Are tag clouds useful for navigation? a network-theoretic analysis. *International Journal of Social Computing and Cyber-Physical Systems*, 1(1):33–55", 2011.

[11] P. Heymann, A. Paepcke, and H. Garcia-Molina. Tagging human knowledge. In *Proceedings of the third ACM international conference on Web search and data mining*, WSDM '10, pages 51–60, New York, NY, USA, 2010. ACM.

[12] J. Kleinberg. The small-world phenomenon: An algorithmic perspective. In *Proceedings of the 32nd ACM Symposium on Theory of Computing*, 2000.

[13] J. M. Kleinberg. Navigation in a small world. *Nature*, 406(6798), August 2000.

[14] J. M. Kleinberg. Small-world phenomena and the dynamics of information. In *Advances in Neural Information Processing Systems (NIPS) 14*, page 2001. MIT Press, 2001.

[15] C. Körner, D. Benz, A. Hotho, M. Strohmaier, and G. Stumme. Stop thinking, start tagging: Tag semantics arise from collaborative verbosity. In *19th International World Wide Web Conference (WWW2010), Raleigh, NC, USA, April 26-30, ACM*, 2010.

[16] J. R. Landis and G. G. Koch. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):pp. 159–174, 1977.

[17] N. Levine. Strongly connected sets in topology. *The American Mathematical Monthly*, 72(10):pp. 1098–1101, 1965.

[18] A. Mehta, A. Saberi, U. Vazirani, and V. Vazirani. Adwords and generalized online matching. *J. ACM*, 54, October 2007.

[19] D. Millen and J. Feinberg. Using social tagging to improve social navigation. In *Workshop on the Social Navigation and Community Based Adaptation Technologies. Dublin, Ireland.* Citeseer, 2006.

[20] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[21] O. Nov and C. Ye. Why do people tag?: motivations for photo tagging. *Commun. ACM*, 53:128–131, July 2010.

[22] J. Sinclair and M. Cardew-Hall. The folksonomy tag cloud: when is it useful? *Journal of Information Science*, 34:15, 2008.

[23] M. Strohmaier, D. Helic, D. Benz, C. Körner, and R. Kern. Evaluation of folksonomy induction algorithms. *ACM Trans. Intell. Syst. Technol.*, 2012.

[24] C. Trattner. Improving the navigability of tagging systems with hierarchically constructed resource lists and tag trails. *CIT*, 19(3):155–167, 2011.

[25] C. Trattner. Linking related content in web encyclopedias with search query tag clouds. *IJWI*, 9(1):33–55, 2011.

[26] C. Trattner, I. Hasani-Mavriqi, D. Helic, and H. Leitner. The austrian way of wiki(pedia)!: development of a structured wiki-based encyclopedia within a local austrian context. In *Proceedings of the 6th International Symposium on Wikis and Open Collaboration*, WikiSym '10, pages 1–10, New York, NY, USA, 2010. ACM.

[27] C. Trattner and D. Helic. Linking related documents: combining tag clouds and search queries. In *Proceedings of the 10th international conference on Web engineering*, ICWE'10, pages 486–489, Berlin, Heidelberg, 2010. Springer-Verlag.

[28] C. Trattner, D. Helic, and S. Maglajlic. Enriching tagging systems with google query tags. In *Information Technology Interfaces (ITI), 2010 32nd International Conference on*, pages 205 –210, june 2010.

6

Navigational Efficiency of Broad vs. Narrow Folksonomies

This chapter is based on the paper *"Navigational efficiency of broad vs. narrow folksonomies"* presented at the 23rd ACM Conference on Hypertext and Social Media in 2012.

It continues the work on the navigational efficiency of tags compared to other tag-alike meta-data structures. In the previous chapter we have shown that search query tag clouds perform extremely well in linking and navigating related content. In this chapter we explore the navigational differences and the similarities of another tag-alike meta-data construct. In particular, we present a study that explores the navigational differences between broad folksonomies which are based on tags provided by users and narrow folksonomies that are grounded on keywords provided by authors. We study both kinds of folksonomies on a dataset provided by Mendeley - a collaborative platform where users can annotate and organize scientific articles with tags and keywords. Our experiments show that tags are more useful navigation than keywords.

The original contribution can be found in [12].

6.1 Abstract

Although many social tagging systems share a common tripartite graph structure, the collaborative processes that are generating these structures can differ significantly. For example, while resources on Delicious are usually tagged by all users who bookmark the web page `cnn.com`, photos on Flickr are usually tagged just by a single user who uploads the photo. In the literature, this distinction has been described as a distinction between *broad vs. narrow folksonomies*. This paper sets out to explore navigational differences between broad (based on tags provided by users) and narrow

(based on keywords provided by authors) folksonomies in social hypertextual systems. We study both kinds of folksonomies on a dataset provided by Mendeley - a collaborative platform where users can annotate and organize scientific articles with tags and keywords. Our experiments suggest that broad folksonomies are more useful navigation than narrow ones.

6.2 Introduction

In social tagging systems, users organize information using so-called *tags* – a set of freely chosen words or concepts – to annotate various resources such as web pages on Delicious, photos on Flickr, or scientific articles on BibSonomy. In addition to using tagging systems for personal organization of information, users can also socially share their annotations with each other. The information structure that emerges through such processes has been typically described as "folksonomies[1]" (**fol**k-generated ta**xonomies**). Usually, such folksonomies are represented as tripartite graphs with hyper edges. These structures contain three finite, disjoint sets which are 1) a set of users $u \in U$, 2) a set of resources $r \in R$ and 3) a set of tags $t \in T$ annotating resources R. A folksonomy as a whole is defined as the annotations $F \subseteq U \times T \times R$ (cf. [28]). A *bookmark* or *post* refers to a single resource r and all corresponding tags t of a user u.

Although this tripartite structure of folksonomies can be mapped onto a broad range of different systems in heterogeneous domains (such as Delicious, Flickr, Mendeley and others), the *collaborative processes that are generating these structures can differ significantly.* For example: While resources on Delicious are usually tagged by a larger group of users (e.g. by everybody who has bookmarked the web page `cnn.com`), photos on Flickr are usually tagged just by a single user (e.g. just by the user who has uploaded the photo). In past discussions, this distinction has been described as a distinction between *broad vs. narrow folksonomies*[2].

Thus, while broad folksonomies are structures that have been generated as a result of aggregating data from *many people tagging the same resource*, narrow folksonomies are structures that have been generated as a result of aggregating data from *single users tagging their own resources.* Although both kinds of folksonomies can be mapped onto the tripartite structure of folksonomies, it is reasonable to expect that they differ with regard to their overall network characteristics and topology, form and function. In this paper we will argue that without thorough investigations of the different characteristics of different kinds of folksonomies (e.g. broad vs. narrow), our understanding of the potentials and limitations of social tagging systems will be limited. Therefore, understanding the usefulness and utility of differ-

[1] `http://www.vanderwal.net/folksonomy.html`
[2] `http://personalinfocloud.com/2005/02/explaining_and_.html`

ent kinds of folksonomies for different tasks - such as navigation, emergent semantics or information retrieval - represents a problem of both theoretical and practical importance.

Similar classifications of meta-data have been analyzed in other application areas such as learning objects meta-data. In their analysis in [31] the authors distinguish between "authoritative" meta-data that is provided by official data descriptors, e.g. learning object authors and "non-authoritative" meta-data which emerges through the usage of learning objects in different contexts, e.g. it is created by a user community. In our terminology "authoritative" meta-data corresponds to narrow folksonomies and "non-authoritative" meta-data to broad folksonomies. The authors argued in their study that there are significant differences in the utility of different types of meta-data. For example, they demonstrated that the "non-authoritative" meta-data is crucial for effective discovery and reuse of learning objects in different contexts.

In this paper, we aim to systematically compare differences between broad and narrow folksonomies on a large tagging system (Mendeley). Mendeley is a collaborative platform for scientists where users can annotate and organize scientific articles with tags. Because Mendeley not only captures data about the set of tags assigned by users, but also about the set of keywords assigned by the authors of articles (extracted from library and meta-data information), *we can generate both broad and narrow folksonomies for the same set of resources* (i.e. scientific articles) at the same time. This means that we can generate broad folksonomies based on the tags users assigned to scientific articles, *and* we can generate narrow folksonomies for the same set of resources based on the keywords that authors assigned to their papers.

In this work, we will compare the usefulness of broad vs. narrow folksonomies for a given *task*: navigation. We start by applying hierarchical clustering algorithms (such as the algorithm by [2] and others) to create hierarchies of tags and keywords as navigational structures between resources. We then use an existing framework for simulating navigation in social tagging systems [15] based on Kleinberg's decentralized search [19] to simulate a hypothetical user navigating the resource space using information provided by keywords vs. tags. In particular, we are going to model a navigational task where the user starts at an arbitrary keyword/tag and navigates to another keyword/tag to reach the list of articles with that keyword/tag. In our simulations, we adopt a greedy routing strategy based on Kleinberg's decentralized search. As a result, we use keyword/tag hierarchies as background knowledge that guides the simulation towards a particular destination by providing information on distances between keywords/tags in the resource network. To reflect the limitations of a real-world user interface, we then repeat the simulations by introducing constraints related to different user interface elements inspired by previous work [13]. The overall outcome of our investigations allows us to shed light on the differences between broad vs.

narrow folksonomies in theoretical but also in practical navigation settings (by considering UI constraints). For our simulations we use a dataset that currently includes about 150 million scientific articles and has a community of about 1,5 million of users who tag articles in an unconstrained manner.

Our results suggest that both broad (tag-based) and narrow (keyword-based) folksonomies support efficient navigation in theory. However, taking some practical limitations of typical user interfaces into account, we find that broad folksonomies outperform narrow folksonomies significantly on our dataset.

In summary, this paper reports on the following findings based on our dataset:

- Narrow folksonomies create less effective navigational structures than broad folksonomies when real-world user interface constraints are considered.
- Our analysis suggests that navigational effectiveness of tags comes from the different viewpoints of readers provided through tagging resources.
- Broad folksonomies provide substantially higher quality of navigational structures than narrow ones. We speculate that with growing numbers of tags in broad folksonomies, their navigational advantage becomes even greater. More research on this question is warranted though.

The remainder of this paper is organized as follows. In Section 6.3, we discuss related work. In Section 6.4 we shortly present our simulation model for user navigation. In Section 6.5, we outline our experimental setup and in Section 6.6 we present our experimental results. In Section 6.7 we discuss the results and provide a possible explanation for the observed difference in navigational efficiency.

6.3 Related Work

Related work in this field of research can be split up into two different parts: *folksonomies*, and *navigation and hierarchies in networks*.

Folksonomies: In the past, folksonomies have been studied from at least two different perspectives – from an ontological and an information retrieval perspective. From the ontological perspective, our community analyzed emergent semantic structures. For example [2, 16, 26] propose algorithms for constructing semantically sound tag hierarchies from social tagging data. A detailed analysis of approaches to semantic relatedness of tags in social tagging systems can be found in e.g. [6]. In our own previous work [22, 23], we investigated the extent to which tag semantics are influenced by user motivation and usage practices. In [33] we investigated the quality of semantic relations in automatically constructed tag hierarchies. By measuring Taxonomic Recall and Precision [9] against a huge number of existing human created concept hierarchies we have shown that algorithms such as e.g. [2] outperform other popular tag hierarchy induction approaches such as Affinity Propagation [11] or Hierarchical K-Means [10].

From the information retrieval perspective, Chi at al. [7] investigated the ability of tags to efficiently encode resources for later retrieval and found out that this ability decreases over time. In [17] and [1] the authors proposed and evaluated search ranking algorithms such as FolkRank and SocialSimilarity Rank. In our own previous work [15], we evaluated the suitability of different tag hierarchies to support navigation in social tagging systems on a theoretical level – not taking user interface constraints into account. There we showed that tag hierarchies created with algorithms such as [2, 16] are able to, at least in theory, provide an efficient support for *navigation* in tagging systems. In subsequent work, we also modeled typical limitations of a standard user interface such as e.g. *directories*, and were able to deduce a new algorithm that produces tag hierarchies that are still able to support efficient navigation even when restricted by a real-world user interface [13]. These hierarchies were evaluated by simulations with the same decentralized approach as it is also used in this paper.

Navigation and hierarchies in networks: Research on navigation in complex networks was initiated by the famous small-world experiment conducted by Milgram [29]. In that experiment randomly selected persons were required to pass a letter to a target person through their social networks. The striking result of the experiment was that the average chain length length was only six. Apart from the findings that humans in that social network are connected by short paths, another conclusion was that humans can efficiently navigate social networks although they have only *local knowledge* of that network – humans can efficiently perform *decentralized search*. Kleinberg concluded that humans possess *background knowledge* of the network structure and that this knowledge allows humans to efficiently find short paths [18, 20, 21]. Kleinberg represented such background knowledge as a hierarchy of nodes, where more similar nodes are situated closer to each other in the hierarchy.

In [32] the authors extend the notion of background knowledge to the notion of *hidden metric spaces*. In such hidden metric spaces nodes are identified by their co-ordinates – distance between nodes is their geometric distance in a particular metric space. Navigation strategies in complex networks are then based on the distances between nodes – an agent always navigates to the node with the smallest distance to a particular destination node. An interesting research question is the structure of such hidden metric spaces that underlie observable networks. In [4], the authors introduce a model with the circle as a hidden metric space and show its effects on routing in the global airport network. In [24] the authors discuss hyperbolic geometry as a hidden metric space (which can be approximated by a node hierarchy) whereas in [5] the authors apply hyperbolic geometry as a model of the hidden metric space of the Internet and design a novel greedy Internet routing algorithm. In [25] the authors describe a novel decentralized search model for efficient navigation in social networks. The model is based on the users interest. By simulating navigation on the co-author network of DBLP[3] they evaluate the model and show the importance of one step lookahead in decentralized search algorithms for social networks.

Hierarchies that are extracted from networks play an important role in many of these network navigation models. Apart from the tag hierarchy induction algorithms based on bipartite networks such as e.g. [16, 2, 13], researchers also proposed hierarchy extraction algorithms for general networks. In [30] the authors discuss an algorithm for hierarchy construction in Wikipedia networks based on metrics for estimating hierarchy level of single nodes. Also, Clauset et al. [8] present a hierarchy induction algorithm based on prediction of hierarchical links. Links prediction problem (in general settings) has been also studied by Liben-Nowell and Kleinberg [27]: They studied the extent to which interactions among members of a social network are likely to occur in the near future.

West and Leskovec [34] performed a study of user navigation behavior. The authors analyzed a collection of click paths of users playing a navigation game in a network of links between the concepts of Wikipedia. In their work they found out that user navigation behavior differs from shortest paths. For example, users typically navigate through high-degree hubs in the early phase and then apply content similarity as a criteria for reaching the destination concept.

6.4 Methodology

Our methodology for comparing the usefulness of broad vs. narrow folksonomies for navigation is simulation. We simulate a user who visits a digital library in search for a set of scientific articles and applies thereby a set of standard information seeking strategies. A recent study that investigated user behavior in Web search [35] showed that not many users satisfy their information need with their first search query. Instead, users visit one of the first search results, follow links on that result page, backtrack, follow some other links, then in many cases refine their search, and so on.

Thus, we model a user who starts the inquiry by issuing a search query either at an external search engine or using the integrated search function provided by the digital library. Upon selecting one of the search results the user lands at a particular page in the digital library and explores the links from that page in order to satisfy her information need. We model this first step by randomly selecting words from broad (tags) vs. narrow (keywords) folksonomies from the library. We represent the user information need as another randomly selected destination keyword together with the list of articles

[3] http://dblp.uni-trier.de/

for which this destination keyword was assigned. We then simulate the navigation from the starting keyword to the destination keyword. In our previous work we simulated the navigation in tagging systems by simulating a user traversing links between tags from tag clouds [14] or links in a hypothetical directory-like user interface for tags [13]. The former was an assessment of the navigability of tags in an unconstrained settings whereas the latter represents a more realistic settings of a user interface that has limitations in the number of items that are presented to the user. Please note that an important advantage of simulation as an evaluation strategy is the possibility to experiment with various configurations and parameters and in this way cover a wide range of different settings – something that would not be possible in more traditional user studies. Thus, we apply the same methodology in this paper and evaluate different settings in which keywords might be used to support navigation, such as unconstrained navigation, or different variations of navigation limited by constraints of a typical user interface.

In [13, 14] we introduced a simple user navigation model – in this paper we just shortly explain its basic principles. Essentially, user navigation in information networks (such as networks of tags, or networks of keywords and scientific articles) is a kind of so-called decentralized search, or search with local knowledge of the network [18, 19, 20, 21]. At each step of navigation towards a specific destination node the user is aware only of links emanating from the current node. The user does not possess the global knowledge of the network and is therefore required to adopt a navigation strategy that will guide her as fast as possible to the destination node. In his research on the search in social networks inspired by the famous small-world experiment by Milgram [29] Kleinberg introduced a simple greedy strategy [18, 19]. The prerequisite for this strategy is the existence of an external background knowledge on the network that defines the notion of distance or similarity between network nodes. An agent applying the greedy strategy selects from currently available links the link that leads to the most similar, i.e. to the node closest to the destination node. Kleinberg was able to show that such a greedy strategy is a very efficient one and that an agent applying that strategy always finds the destination node in a small number of steps that is bounded poly-logarithmically in the number of nodes.

Thus, we simulate user navigation by applying such a greedy strategy in search from the start to the destination node. In [13, 14] we represented the background knowledge as various tag hierarchies. Clearly, the structure of this hierarchy influences navigational capability. We assessed navigational efficiency provided by those hierarchies by measuring how often the search for the destination node is successful and if successful how fast is it. We were able to show in those papers that tag hierarchies can indeed support efficient navigation. We also designed a new algorithm that induces tag hierarchies that are efficiently navigable even under the restrictions of a realistic user interface. In this paper we apply those same algorithms on collections of keywords and scientific articles, measure the navigability of keywords and compare those results with the results that we obtained for tags on the same set of resources.

Moreover, in this paper we extend our navigation model to account for a situation where the user looks for a specific scientific article. Thus, we are not only interested in how quickly we can find keywords – we also want to know how easy it is to find a particular article once when we reach one of its keywords.

6.5 Experimental Setup

6.5.1 Simulation and Evaluation Metrics

We divide our evaluation into two parts: We compare the usefulness of broad vs. narrow folksonomies by comparing their (i) encoding efficiency and (ii) navigational efficiency.

Encoding efficiency. First, we evaluate how good different folksonomical data is at encoding articles for later retrieval. This evaluation provides an insight in the intermediate exploration steps of the navigation process – the user has already reached a potentially interesting keyword or tag and the system presents a list of articles associated with that keyword or tag. We want to estimate how easy is it to find a specific article in this list. This is typically measured in terms of conditional entropy [7]. Entropy is a measure of uncertainty in a random variable. In information theory entropy is expressed in the number of bits that are needed to encode a random variable. Entropy reaches the maximal value when the random variable is distributed uniformly (uncertainty in the value of that random variable is maximal) and is minimal, i.e. it is equal to zero if the random variable always takes on a single value. Entropy of a single random variable (e.g. tags or keywords) is calculated by:

$$H(X) = -\sum_{x \in X} p(x) log(p(x)) \tag{6.1}$$

In turn, conditional entropy quantifies uncertainty in one random variable (articles) once we know a specific value of another random variable (keywords or tags). Thus, conditional entropy of articles measures how difficult is to find a specific article within the presented list. Higher values of conditional entropy mean that there is more uncertainty and it is therefore more difficult to reach a particular article. On the contrary, lower values of conditional entropy mean that the first random variable (keywords or tags) encodes articles more efficiently, decrease uncertainty, and thus it is easier for users to reach a specific article. Conditional entropy of two random variables is given by:

$$H(Y|X) = -\sum_{x \in X} p(x) \sum_{y \in Y} p(y|x) log(p(y|x)) \qquad (6.2)$$

The navigability evaluation consists of four steps:

Network construction. We start with the datasets that include triples of keywords or tags, articles, and authors or users. From those datasets we construct bipartite networks of keywords (tags) and articles and remove the user information as that information is typically not relevant for navigation. Subsequently, we project the bipartite networks onto keyword-to-keyword and tag-to-tag networks as those networks are available for the user for navigation. We assume that article lists are also presented to the user upon selecting a keyword or a tag but only as a means of satisfying the initial information need, whereas keywords or tags are used for exploration, i.e. as a means of making progress towards the final destination.

Hierarchy construction. We induce broad (tag-based) and narrow (keyword-based) folksonomy hierarchies which we will use as the background knowledge to steer navigation. We use two algorithms for constructing hierarchies. In [16], the authors introduce a generic algorithm for producing hierarchies from bipartite networks such as tag-to-resource networks. The algorithm can be applied to arbitrary bipartite structures. The algorithm takes as input two parameters. The first is a ranked list of tags sorted by their centrality in the projected tag-to-tag network. This centrality ranking acts as a proxy to the generality ranking of tags. Benz et al. [3] showed that the centrality provides a viable approximation to the tag generalities. The second input parameter is the tag similarity matrix. The algorithm starts then by a single node hierarchy with the most general tag as the root node and then iterates through the centrality list. At each iteration step, the algorithm adds the current tag to the hierarchy as a child to its most similar tag. The centrality and similarity measure are exchangeable – in [16] the authors use closeness centrality and cosine similarity, whereas in [2] the authors select degree centrality and co-occurrence similarity measure. As both combinations perform similarly in supporting navigation [14], we select the latter combination because of better computational properties. This algorithm produces unbalanced hierarchies that are typically very broad in the top hierarchy levels. As some of the top nodes in real datasets might end up with hundreds or even thousands of children those hierarchies give us the insight in the intrinsic, theoretical, and unconstrained navigational support. We obtain a more realistic assessment of navigational efficiency by applying a variant of this algorithm. In [13], we extended that algorithm and introduced an algorithm that takes also the branching factor (the maximal number of children) of the final hierarchy as an input parameter. Through re-balancing of the hierarchy and introduction of nested misc categories we were able to produce hierarchies that support efficient navigation even under realistic limitations imposed by a typical user interface.

Search pairs selection. We randomly select one million of so-called search pairs consisting of a start node and a destination node. Both start and destination nodes are low degree nodes as searching for high degree nodes is a trivial task. For those pairs we calculate the global shortest path that we will use as our metric to assess the navigation efficiency.

Navigation simulation. We run simulation with greedy navigation on those search pairs and measure the success rate s and stretch τ which is the ratio of the number of simulator steps and the global shortest path. We calculate the global averages of both metrics (s_g and τ_g), as well as distribution of both values over the global shortest path. Also we calculate average of the global shortest path ($\bar{l}$), as well as average number of simulator hops ($\bar{h}$), i.e. average number of clicks of the simulated user.

6.5.2 Datasets

Mendeley[4] claims to be the largest research database, with 150 million papers and 1,5 million users. For our experiments, we used tagging data (dataset **T**) from the system gathered in September 2011 as well as a snapshot from the Mendeley system which includes papers as well as the corresponding keywords provided by the authors (dataset **K**). For dataset **T** we lowercased the tags and removed typos and personal bookmarks, i.e. tags that were used only once by a single user. Lowercasing of the keywords was also performed for dataset **K**. Furthermore we constructed an "overlapped" dataset - a dataset which includes only articles for which both keywords and tags are available. These datasets are called **OT** – overlapped tags and **OK** – overlapped keywords respectively.

Projection of the Dataset: After this preprocessing step, we construct the bipartite networks of keywords and articles and tags and articles. From those bipartite networks we extract the largest connected component (which typically contains around 99% of the network nodes). Finally, we project the largest connected component onto keyword-to-keyword and tag-to-tag networks and obtain the final networks on which we perform our analysis. The dataset and network statistics are shown in Table 6.1.

The first important property here to note is that the quantitative ratio of the number of links and the number of meta-data items (i.e. nodes) is comparable between the data set. The second property – the effective diameter (which is the longest shortest path that connects 90% of all network nodes) – is also comparable in all datasets. Thus, this basic *quantitative network-theoretic properties* indicate that all networks possess similar navigational properties. Hence, any differences in navigational efficiency have to be accounted for *qualitative differences in the network topology.*

[4] http://www.mendeley.com

	K	T	OK	OT
Bipartite				
meta-data	1,124,260	399,703	469,952	201,651
Links	28,459,841	12,869,137	3,323,787	1,492,217
Articles	5,172,180	3,649,350	523,488	523,488
$\frac{\#Links}{\#meta-data}$	25.3	32.3	140.8	134.72
Eff.Diam	6.92	7.10	8.25	8.65
Projected Dataset				
meta-data	1,092,655	371,044	455,001	166,957
Links	124,690,988	47,760,792	26,450,686	7,877,564
$\frac{\#Links}{\#meta-data}$	114.18	128.7	58.13	47.5
Eff.Diam	4.06	3.94	4.79	4.68

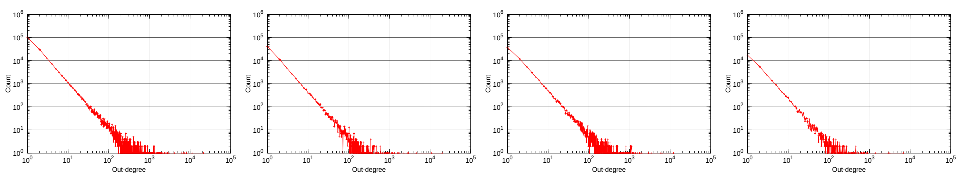
Table 6.1: Dataset and network statistics of broad (T, OT) vs. narrow (K, OK) folksonomies. Datasets OT and OK only contain articles for which both tags and keywords are available.

	K	T	OK	OT
Entropy	15.09	14.23	12.74	12.39
Cond. Entropy	6.40	5.92	4.18	3.81

Table 6.2: Entropy and Conditional Entropy for broad (T, OT) vs. narrow (K, OK) folksonomies. Datasets OT and OK only contain articles for which both tags and keywords are available.

6.6 Results

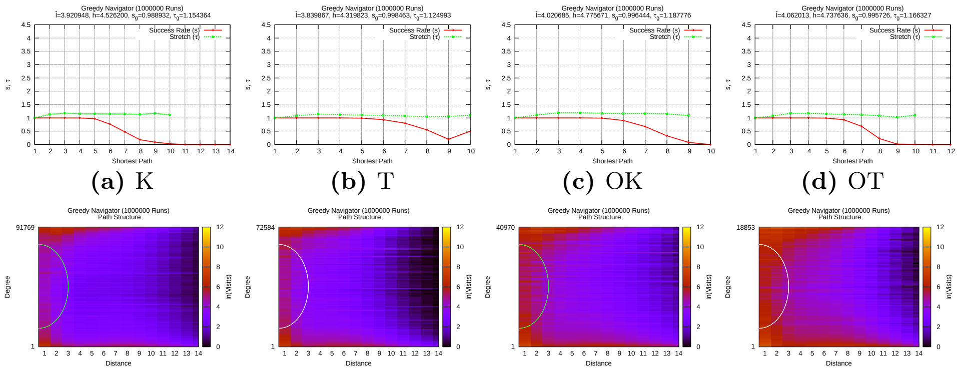
6.6.1 Tag and Keyword Entropy



(a) K (b) T (c) OK (d) OT

Figure 6.1: Out degree distribution of unconstrained hierarchies. The top hierarchy levels are populated by high-degree hubs – nodes that have hundreds or even thousands of children nodes. The hierarchies are very broad and flat.

Table 6.2 shows the entropy of articles conditional on keywords and tags in all four datasets. Although it is difficult to interpret absolute values obtained for the conditional entropy, a comparison of entropy values obtained for different datasets provides insight in the relative encoding efficiencies of

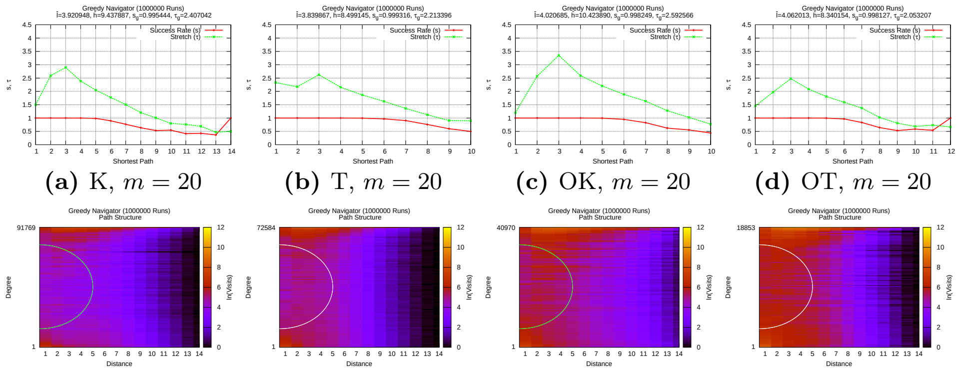


(a) K **(b)** T **(c)** OK **(d)** OT

(e) K **(f)** T **(g)** OK **(h)** OT

Figure 6.2: Results of the simulation with unconstrained user interface. *Top:* Average shortest path $\bar{l}$, average hop count $\bar{h}$, greedy navigator success rate s and stretch τ – global average values (s_g and τ_g) and distribution over shortest paths. Theoretical evaluation of Mendeley tag hierarchies produces results comparable to other tagging datasets. In theory, tag hierarchies support efficient navigation – both success rate and stretch are close to 1. Similarly, keyword hierarchies aid efficient navigation – success rate and stretch are excellent. *Bottom:* Navigator path structure without user interface constraints. The density maps visualize visit frequency to nodes of a given degree at a given distance to the destination node – the color is logarithm of the visit frequency (black and violet indicating less visits; orange and yellow indicating more visits). Over all datasets, the top nodes are the most visited nodes – these are the nodes from the network core where the phase transition in the navigation process occurs. A specific property of navigation paths in tagging networks are so-called shortcuts between related mid-degree nodes occurring at the smaller distances to the destination node (see e.g. white marked region of a large orange-colored area in 6.2h). Those shortcuts are taken between sibling tags of a high-degree parent in the cases where the destination node is situated in the sub-hierarchy of one of the siblings. The density maps reveal a slightly different path structures between keyword and tag navigation. The green marked regions of shortcut areas in the keyword navigation (6.2e and 6.2g) show that shortcuts between related mid-degree and siblings are taken less frequently in the case of keyword navigation – high-degree hubs are more frequently visited in keyword than in tag networks. Since the global success rate and stretch in both networks are comparable to each other this phenomenon indicates that there exist structural differences between keyword and tag hierarchies – a possible explanation would be that tag hierarchies are somewhat richer in structure, i.e. keyword hierarchies more broad and flat.

broad vs. narrow folksonomies. From this analysis we can observe that in our dataset, broad folksonomies (T, OT) encode articles more efficiently than narrow folksonomies (K, OK). In other words, on average we know more about articles annotated by a particular tag than about articles annotated by a particular keyword. This is important when considering that users



(a) K, $m = 20$ (b) T, $m = 20$ (c) OK, $m = 20$ (d) OT, $m = 20$

(e) K, $m = 20$ (f) T, $m = 20$ (g) OK, $m = 20$ (h) OT, $m = 20$

Figure 6.3: Results of the simulation with constrained user interface. The number of siblings is limited to $m = 20$. *Top:* Average shortest path $\bar{l}$, average hop count $\bar{h}$, greedy navigator success rate s and stretch τ – global average values (s_g and τ_g) and distribution over shortest paths. Although the success rates remain excellent over all datasets, stretch increases slightly in keyword datasets. This results in path lengths that are on average longer by 1 or 2 in keyword networks. *Bottom:* Path structure with user interface constraints. The green marked regions of shortcut areas in keyword networks (6.3e and 6.3g) demonstrate less frequent shortcuts than in tag networks (white regions in 6.3f and 6.3h) explaining the increased stretch values in keyword networks.

navigate for resources, not for tags. Our simulation currently does not take into account that users would have to investigate all resources attached to a particular keyword. Hence, the more uncertainty there is on the articles captured by a node, the more time users have to invest for searching the list of articles.

6.6.2 Unconstrained Navigation

We start our navigational analysis with an estimation of the theoretical navigability of keyword and tag hierarchies. Thus, we construct hierarchies by using Heymann's algorithm [16] which does not consider any user interface constraints. The algorithm produces broad and flat hierarchies in which the nodes from the top hierarchies have hundreds or even thousands of children nodes. Figure 6.1 shows the degree distributions of the hierarchies depicting the existence of hub nodes.

The results of the simulation with Mendeley tags are comparable with our previous experiments with tagging datasets from Flickr, Delicious, LastFM, BibSonomy, and CiteULike [13, 14]. In such theoretical settings Mendeley tags are efficiently navigable. Keyword networks show similar behavior – in theory, keywords support efficient navigation. The complete results of the experiments are shown in Figure 6.2.

6.6.3 Constrained Navigation

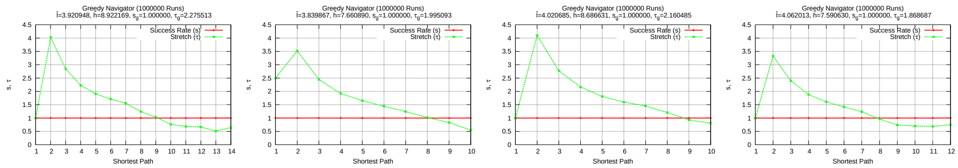
In our next experiments we configure the simulator to reflect typical limitations of a standard user interface, e.g. a directory-like interface, such as Yahoo directory[5]. Thus, we model constraints such as limited number of children nodes that are shown (e.g. 20 children), limited number or related items (e.g. 20 siblings), or combination of both restrictions. As we have shown in [13], such restrictions seriously impede the navigation properties of tag hierarchies and we obtain similar results for both keyword and tag hierarchies. The most interesting observation that we make with those experiments is the difference in stretch values for the limitation of the number of related items that are presented to the user. In our experiments, we observe increased stretch values for keyword navigation resulting in one or two more clicks that are needed on average to reach the destination node. This result is consistent over all datasets and it might reflect an intrinsic property of keyword networks and keyword hierarchies. Our explanation for this phenomenon is that within a group of co-occurring keywords there exist a single keyword which "dominates" the group, i.e. other keywords co-occur more frequently with that "dominating" keyword and less frequently with other keywords from the group. The "dominator" becomes a parent node in the hierarchy and all other nodes are attached as children to that node (see also 6.7). Thus, the limitation of the number of siblings that are presented to the user causes that a longer path over the parent node is taken and increases the path length by 1 or 2 (see Figure 6.3).

6.6.4 Realistic Constrained Navigation

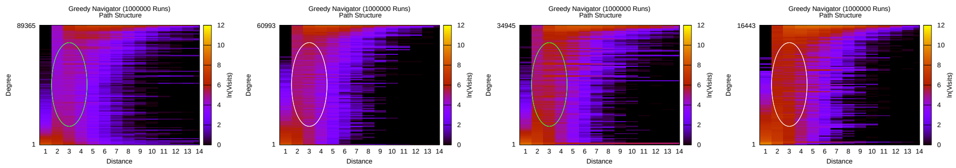
Finally, we want to perform experiments using an alternative algorithm for hierarchy induction to better reflect the realities of user interfaces. We apply the algorithm presented in [13] that produces balanced hierarchies with a maximal number of children (we set e.g. 20 children to reflect a typical user interface limitation). The algorithm produces a nested sub-hierarchy of so-called misc categories in which it inserts nodes with the smallest similarities to their parent node. In a typical case, low-degree nodes from the long tail are inserted into such nested misc categories. In our experiments, we obtain similar results as in experiments limiting the number of siblings. Consistently and over all datasets, keywords perform slightly worse exhibiting increased stretch and an increase of the average number of clicks by 1 (see Figure 6.4).

Finally, we remove misc categories completely to reflect the situation where users might not navigate within these categories. In those experiments we obtain smaller success rates that are comparable to each other over all datasets.

[5] http://dir.yahoo.com/



(a) K (b) T (c) OK (d) OT



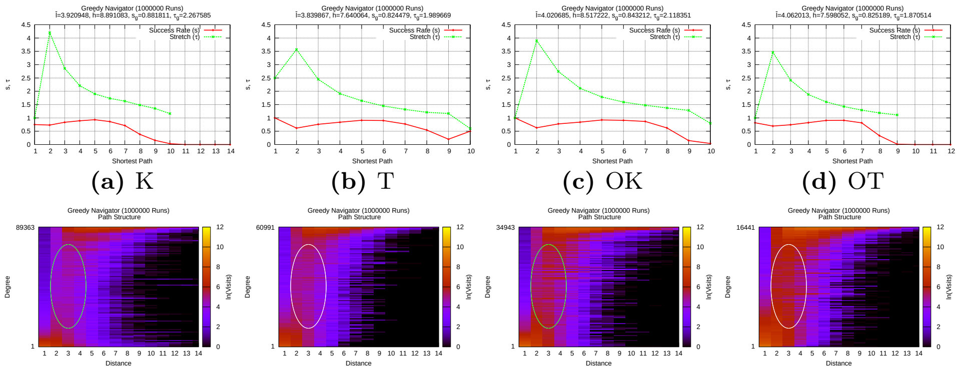
(e) K (f) T (g) OK (h) OT

Figure 6.4: Results of the simulation with balanced hierarchies. The number of children and siblings is set to 20. *Top:* Average shortest path $\bar{l}$, average hop count $\bar{h}$, greedy navigator success rate s and stretch τ – global average values (s_g and τ_g) and distribution over shortest paths. As previously observed the success rates remain stable and excellent over all datasets, whereas stretch increases slightly in keyword datasets. This results in path lengths that are on average longer by 1 in keyword networks. *Bottom:* Navigator path structure with balanced hierarchies. Again, the green marked regions of shortcut areas in keyword navigation (6.4e and 6.4g) indicate smaller shortcut frequencies than in tag navigation (white ellipses in 6.4f and 6.4h).

As before, we observe an increased stretch in keyword networks resulting in the average number of clicks to increase by 1 in those networks (see Figure 6.5).

6.7 Discussion

Our results show that in realistic navigational settings - when we take into account user interface limitations - tag navigation is slightly more efficient than keyword navigation. Moreover, tag encoding efficiency is also higher than keyword encoding efficiency. The density maps reveal the reason for this finding – there are more shortcuts taken between mid-degree and high-degree siblings in tag hierarchies than between such keywords in keyword hierarchies. A possible cause for that is a lower average overlap between sibling keywords compared to sibling tags. We will explain this situation with the following simple example. Suppose we have an article dealing with navigation in tagging systems. The authors define the following keywords for this article: "`folksonomy`", "`tagging`", "`navigation`" (see $r1$ in Figure 6.6). Suppose also that the authors calculate entropy in that article, but do not include "`entropy`" as a keyword in their article. Thus, the authors define their single viewpoint that defines a narrow navigation structure in



(a) K (b) T (c) OK (d) OT

(e) K (f) T (g) OK (h) OT

Figure 6.5: Results of the simulation with balanced hierarchies without misc categories. The number of children and siblings is set to 20. *Top:* Average shortest path $\bar{l}$, average hop count $\bar{h}$, greedy navigator success rate s and stretch τ – global average values (s_g and τ_g) and distribution over shortest paths. The success rates is smaller than before over all datasets. Again, stretch increases slightly in keyword datasets. *Bottom:* Navigator path structure with balanced hierarchies. The green marked regions of shortcut areas in the keyword navigation (6.5e and 6.5g) and white marked regions in tag datasets (6.5f and 6.5h) show differences in the number of shortcuts.

the proximity of that article and its keywords. Now, suppose that multiple users annotate that article with tags. For example, the first user annotates it with "`folksonomy`" and "`tagging`". The second user annotates it with "`navigation`", and the third user with "`entropy`" (because that is the most interesting part of the article for that user). Now, there are multiple viewpoints on the same article and there are multiple navigational structures that are broader and *overlap* with each other. Suppose now that a user is interested in an article about entropy. Now a user may reach that article in a number of alternative ways – one of these paths leads also over our sample article as its "`entropy`" tag represents an entrance to a completely different cluster in the network. Thus, the user might come from a cluster related to e.g. social tagging and then upon arriving on the sample article take a *shortcut* over "`entropy`" tag and enter the entropy cluster. Thus, tags provide different, alternative, and more heterogeneous access paths to articles. In other words, tag folksonomies result in tag distributions whereas keyword folksonomies result in simple almost independent groups of keywords.

Moreover, such multiple viewpoints from many users tagging the same resource collection result in richer hierarchical structures – at least under the algorithms that we applied in our paper. Figure 6.6 depicts an example with a group of similar articles dealing with e.g. social tagging systems – the constructed hierarchies differ in their structures. Richer structures

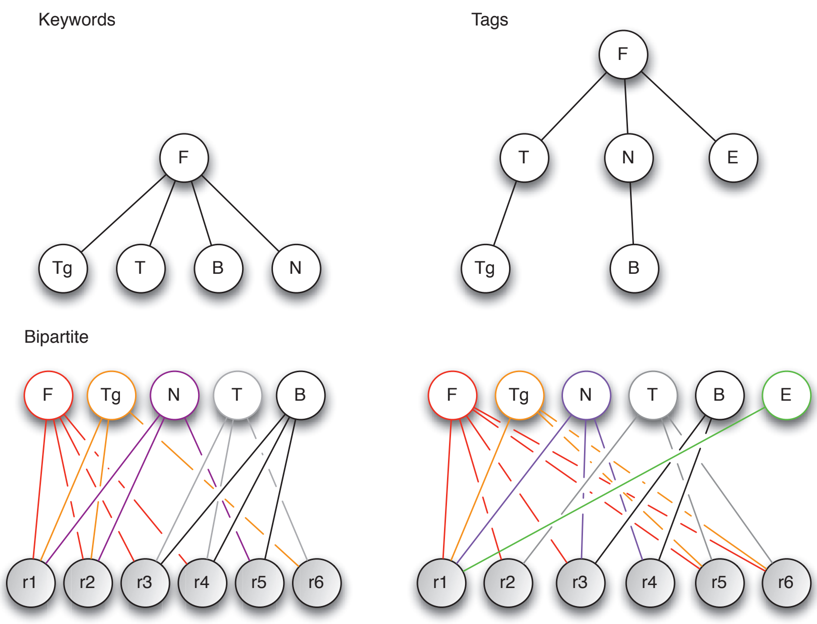


Figure 6.6: Two simple examples showing the emergence of hierarchies in keyword networks (left) and tag networks (right) with meta-data "`folksonomy`" (**F**), "`tagging`" (**Tg**), "`tags`" (**T**), "`navigation`" (**N**), "`browsing`" (**B**), and "`entropy`" (**E**). In keyword (narrow) folksonomies keywords are applied for grouping of articles. On contrary, in tag (broad) folksonomies tags are assigned by many users with multiple and possible alternative viewpoints. This results in tag distributions that impose richer overlap between similar tags. As a consequence the hierarchies that are based on tag generality and their mutual similarities are richer in structure than keyword hierarchies. Our experiments show that those structurally richer hierarchies are more stable and robust to the negative effects of the user interface constraints.

that emerge in tag hierarchies are more robust to the restrictions imposed by a user interface – less tags are affected by e.g. limiting the number of related tags as compared to more keywords that are removed when we limit the number of related keywords presented to the user. We can provide a remedy for this problem of keyword networks by e.g. enriching the keywords with additional meta-data such as categorizations, or subject descriptors to turn *narrow* keyword folksonomy into a *broad* folksonomy similar to the tag folksonomy.

6.8 Conclusions

This paper set out to study differences between broad vs. narrow folksonomies and their usefulness for the task of navigation. Using data from Mendeley, we created both broad (based on tags provided by users) and narrow (based on keywords provided by authors) folksonomies. While our experiments show that broad and narrow folksonomies exhibit comparable quantitative properties, we find interesting qualitative differences with regard to navigation. For example, broad folksonomies create more efficient navigational structures that enable users to find target resources with fewer hops. We find that the reason for better navigational utility of broad folksonomies can be explained by the fact that greater overlap between tags provides better options for users to switch between different parts of the network. Narrow folksonomies are not able to provide this kind of support. While our findings are limited to a single dataset (Mendeley), they warrant future research in this direction. Our results are relevant for designers of social tagging systems and for engineers aiming to improve the navigability of their systems.

References

[1] S. Bao, G. Xue, X. Wu, Y. Yu, B. Fei, and Z. Su. Optimizing web search using social annotations. In *Proceedings of the 16th international conference on World Wide Web - WWW '07*, page 501, New York, New York, USA, May 2007. ACM Press.

[2] D. Benz, A. Hotho, and G. Stumme. Semantics made by you and me: Self-emerging ontologies can capture the diversity of shared knowledge. In *Proc. of the 2nd Web Science Conference (WebSci10)*, Raleigh, NC, USA, 2010.

[3] D. Benz, C. Körner, A. Hotho, G. Stumme, and M. Strohmaier. One tag to bind them all : Measuring term abstractness in social metadata. In *Proceedings of the 8th Extended Semantic Web Conference (ESWC 2011)*, Heraklion, Crete, May 2011.

[4] M. Boguñá, D. Krioukov, and K. C. Claffy. Navigability of complex networks. *Nature Physics*, 5:74–80, Jan. 2009.

[5] M. Boguñá, F. Papadopoulos, and D. Krioukov. Sustaining the Internet with hyperbolic mapping. *Nature Communications*, 1:62, Sept. 2010.

[6] C. Cattuto, D. Benz, A. Hotho, and G. Stumme. Semantic grounding of tag relatedness in social bookmarking systems. In *Proc. of International Semantic Web Conference 2008*, volume 5318 of *LNAI*, pages 615–631, Heidelberg, 2008. Springer.

[7] E. H. Chi and T. Mytkowicz. Understanding the efficiency of social tagging systems using information theory. In *HT '08: Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*, pages 81–88, New York, NY, USA, 2008. ACM.

[8] A. Clauset, C. Moore, and M. E. J. Newman. Hierarchical structure and the prediction of missing links in networks. *Nature*, 453(7191):98–101, 2008.

[9] K. Dellschaft and S. Staab. On how to perform a gold standard based evaluation of ontology learning. In *Proceedings of ISWC-2006 International Semantic Web Conference*, Athens, GA, USA, November 2006. Springer, LNCS.

[10] I. Dhillon, J. Fan, and Y. Guan. Efficient clustering of very large document collections. In R. Grossman, C. Kamath, and R. Naburu, editors, *Data Mining for Scientific and Engineering Applications*. Kluwer Academic Publishers, 2001.

[11] B. J. J. Frey and D. Dueck. Clustering by passing messages between data points. *Science*, 315(5814):972–976, January 2007.

[12] D. Helic, , C. Körner, M. Granitzer, M. Strohmaier, and C. Trattner. Navigational efficiency of broad vs. narrow folksonomies. In *Proceedings of the 23rd ACM Conference on Hypertext and Social Media*, HT '12, New York, NY, USA, 2012. ACM.

[13] D. Helic and M. Strohmaier. Building directories for social tagging systems. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, CIKM '11, pages 525–534, New York, NY, USA, 2011. ACM.

[14] D. Helic, M. Strohmaier, C. Trattner, M. Muhr, and K. Lermann. Pragmatic evaluation of folksonomies. In *Proc. of the 21st International World Wide Web conference, WWW '11*, New York, NY, USA, 2011. ACM.

[15] D. Helic, C. Trattner, M. Strohmaier, and K. Andrews. Are tag clouds useful for navigation? a network-theoretic analysis. *International Journal of Social Computing and Cyber-Physical Systems*, 1(1):33–55", 2011.

[16] P. Heymann and H. Garcia-Molina. Collaborative creation of communal hierarchical taxonomies in social tagging systems. Technical Report 2006-10, Stanford InfoLab, April 2006.

[17] A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. Folkrank: A ranking algorithm for folksonomies. In *Proc. FGIR 2006*, 2006.

[18] J. Kleinberg. The small-world phenomenon: An algorithmic perspective. In *Proceedings of the 32nd ACM Symposium on Theory of Computing*, 2000.

[19] J. Kleinberg. Complex networks and decentralized search algorithms. In *International Congress of Mathematicians (ICM)*, pages 1019–1044, Zürich, Switzerland, 2006. European Mathematical Society Publishing House.

[20] J. M. Kleinberg. Navigation in a small world. *Nature*, 406(6798), August 2000.

[21] J. M. Kleinberg. Small-world phenomena and the dynamics of information. In *Advances in Neural Information Processing Systems (NIPS) 14*, page 2001. MIT Press, 2001.

[22] C. Körner, D. Benz, A. Hotho, M. Strohmaier, and G. Stumme. Stop thinking, start tagging: Tag semantics arise from collaborative verbosity. In *19th International World Wide Web Conference (WWW2010), Raleigh, NC, USA, April 26-30, ACM*, 2010.

[23] C. Körner, R. Kern, H. P. Grahsl, and M. Strohmaier. Of categorizers and describers: An evaluation of quantitative measures for tagging motivation. In *21st ACM SIGWEB Conference on Hypertext and Hypermedia (HT 2010), Toronto, Canada, ACM*, Toronto, Canada, June 2010.

[24] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá. Hyperbolic geometry of complex networks. *Phys. Rev. E*, 82(3):036106, Sep 2010.

[25] S. Lattanzi, A. Panconesi, and D. Sivakumar. Milgram-routing in social networks. In *Proceedings of the 20th international conference on World wide web*, WWW '11, pages 725–734, New York, NY, USA, 2011. ACM.

[26] R. Li, S. Bao, Y. Yu, B. Fei, and Z. Su. Towards effective browsing of large scale social annotations. *Proceedings of the 16th international conference on World Wide Web*, page 952, 2007.

[27] D. Liben-Nowell and J. Kleinberg. The link prediction problem for social networks. In *Proceedings of the twelfth international conference on Information and knowledge management*, CIKM '03, pages 556–559, New York, NY, USA, 2003. ACM.

[28] P. Mika. Ontologies are us: A unified model of social networks and semantics. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(1):5–15, 2007.

[29] S. Milgram. The small world problem. *Psychology Today*, 1:60–67, 1967.

[30] L. Muchnik, R. Itzhack, S. Solomon, and Y. Louzoun. Self-emergence of knowledge trees: Extraction of the wikipedia hierarchies. *Phys. Rev. E*, 76:016106, Jul 2007.

[31] M. Recker and D. Wiley. A non-authoritative educational metadata ontology for filtering and recommending learning objects. *Journal of Interactive Learning Environments*, 9(3):255–271, 2003.

[32] M. A. Serrano, D. Krioukov, and M. Boguñá. Self-similarity of complex networks and hidden metric spaces. *Phys. Rev. Lett.*, 100(7):078701, Feb 2008.

[33] M. Strohmaier, D. Helic, D. Benz, C. Körner, and R. Kern. Evaluation of folksonomy induction algorithms. *ACM Trans. Intell. Syst. Technol.*, 2012.

[34] R. West and J. Leskovec. Human wayfinding in information networks. In *Proceedings of the 21th international conference companion on World wide web*, WWW'12, New York, NY, USA, 2012. ACM.

[35] R. W. White and S. M. Drucker. Investigating behavioral variability in web search. In *Proceedings of the 16th international conference on World Wide Web*, WWW '07, pages 21–30, New York, NY, USA, 2007. ACM.

Part IV

Solution: Build Efficiently Navigable Tag-Based Browsing Constructs

To what extent can we build better tag-based browsing constructs that support efficient search/navigation in tagging systems?

7

On the Construction of Efficiently Navigable Tag Clouds Using Knowledge from Structured Web Content

This chapter is based on the article *"On the Construction of Efficiently Navigable Tag Clouds Using Knowledge From Structured Web Content."* published in the Journal of Universal Computer Science in 2011 and the paper *"Improving the Navigability of Tagging Systems with Hierarchically Constructed Resource Lists: A Comparative Study"* presented at the 33rd International Conference on Information Technology Interfaces in 2011.

It is the first out of three chapters that deals with the question to what extent better tag-based browsing constructs can be created that support efficient navigation in tagging systems. To that end, we present in this chapter a novel tag cloud construction algorithm that uses knowledge from structured Web content to create tag cloud networks which are more efficiently navigable than the one generated by the currently most popular tag cloud construction approach. Contrary to previous work (see Chapter 3), the proposed algorithm takes the semantic relations of the tagging system into account. In a number of experiments based on simulations and a user study, we show the high performance of our approach.

The original contributions can be found in [30] and [28].

7.1 Abstract

In this paper we present an approach to improving navigability of a hierarchically structured Web content. The approach is based on an integration of a tagging module and adoption of tag clouds as a navigational aid for such content. The main idea of this approach is to apply tagging for the purpose of a better highlighting of cross–references between information items across the hierarchy. Although in principle tag clouds have the potential to support

efficient navigation in tagging systems, recent research identified a number of limitations. In particular, applying tag clouds within pragmatic limits of a typical user interface leads to poor navigational performance as tag clouds are vulnerable to the so-called pagination effect. In this paper, solutions to the pagination problem are discussed, implemented as a part of an Austrian online encyclopedia called the Austria-Forum, and analyzed. In addition, a simulation- and user based evaluation of the new algorithm have been conducted. The first evaluation results are quite promising, and show that the proposed algorithm creates tag-network structures which are more navigable than current state-of-the-art approaches for tag cloud construction.

7.2 Introduction

An example of a semi-structured website is the Austria-Forum[1]. Basically, the Austria-Forum is a collection of several hierarchically structured Austrian encyclopedias that contain information about biographies, post stamps, coins, or the Austrian Universal Encyclopedia AEIOU[2]. the Austria-Forum is a Wiki based system, whose articles within a single encyclopedia are hierarchically structured. Thus, the Austria-Forum is also called a structured Wiki [31]. Currently, as of 1^{st} of October 2010 the system provides over 130,000 information items to the user.

Due to the hierarchical structure and the rapid growth of the system over the past few months, links between articles in different encyclopedias are sparse even though they might be related to each other. For example, there are several "Mozart" stamps in the Stamps encyclopedia. However, none of these articles has links to the "Mozart" biography, or "Mozart" coins because the articles are created and managed independently.

To tackle the problem of poor connectivity, a simple tagging mechanism was introduced to the Austria-Forum [32]. In tagging systems people use free-form vocabulary [8] to annotate resources with "tags" [34, 19]. This is either done for semantic reasons (e.g. to enrich information items with metadata), conversational (e.g. for social signaling) [2] or for organizational reasons (e.g. to categorize information items) [16]. Regardless of "why people tag" [27, 22, 17], tags can be visualized in so-called "tag clouds". A tag cloud is a selection of tags related to a particular resource. Upon clicking on a tag, a list of resources tagged with that tag is presented to users leaving them with a possibility to easily navigate to related resources. The main idea of including a tag module into the Austria-Forum can best be described via the previously mentioned "Mozart" example. Suppose that users tag "Mozart" stamps, "Mozart" coins, "Mozart" biography, or any other document dealing with "Mozart" with a common tag, e.g. "Amadeus". Whenever users

[1] http://www.austria-lexikon.at/
[2] http://www.aeiou.at/

navigate to any of these articles a tag cloud containing all assigned tags is presented by the system. Thus, users can now click on "Amadeus" tag and this presents a list of all other articles tagged by that tag. Consequently, all articles tagged with "Amadeus" are now linked to each other, in fact, they are cross-linked across the hierarchical structure. Due to such indirect linking capabilities, tag clouds are often applied to provide navigational support in tagging systems (cf. systems such as Flickr, Delicious, or BibSonomy).

Recently, in a number of studies tag clouds have been investigated from user interface [25, 26, 9] perspectives. These studies agree with regard to some interesting findings, such as the observation that current tag cloud calculation algorithms are not always so useful as one might think. Furthermore, we found that the ability of tag clouds to support "efficient" navigation under the consideration of pragmatic user interface limits, such as tag cloud size and pagination, is very poor [11]. In particular, the pagination effect causes the fragmentation of the network destroying the connected component and thus leaving a majority of resources unreachable.

In this paper, we present an approach to constructing tag clouds that support more efficient navigation than currently available approaches This new algorithm is based on the idea of hierarchical network models [14]. The algorithm has been implemented into the Austria-Forum as a general tool for improving connectivity and to support better navigation of the system as a whole.

The paper is structured as follows: Section 7.3 presents a model for tag cloud based navigation. Section 7.4 discusses the problems of tag cloud based navigation and current tag cloud construction algorithms. Section 7.5 presents the idea of a new and optimized tag cloud calculation algorithm based on the ideas of a hierarchical network model within an online encyclopedia system called the Austria-Forum. Section 7.6 provides an analysis of the potentials and limitation of this new approach. Section 7.7 gives some insights to related work in this field. Finally, Section 7.8 concludes the paper and provides an outlook for the future work in this area.

7.3 Model of Tag Cloud Navigation

In this paper, the tagging data is modeled as a pair of the form (r, t), where r is a resource from the set of all resources R, and t is a tag of all tags T. Here, we do not take into account users as we concentrate only on links between resources imposed by tags assigned to those resources. The main navigational aid in a tagging system is a tag cloud and we denote it with TC. Formally, a tag cloud TC is a particular selection of tags from the tag set.

Due to user interface restrictions the number of tags within a tag cloud is usually limited to an upper bound. To model this situation we additionally introduce a factor n as a maximum number of tags in a tag cloud.

Usually, the most popular tags are assigned to a large number of resources – hundreds or even thousands of resources. When a user clicks on such a tag, tagging systems present a long paginated list of tagged resources. In most cases, 10–100 resources are presented to the users at once (see e.g. Delicious or Bibsonomy). To model these user interface limitation – that we refer to as the pagination from here on – we introduce a factor k that k-limits the resource list of tags within a tag cloud TC.

Finally, let us model the navigation process in a tagging system. Navigation in a tagging system might start from a home page where a system-global tag cloud is presented. Typically, tags with the highest global frequency are selected for inclusion in a tag cloud. Upon clicking on a particular tag a k-limited list of resources is shown. Once the user has selected a specific resource, the system transfers the user to the selected resource and presents a resource-specific tag cloud TC_r. The tags in such a resource-specific tag are selected according to the highest local frequency. In the next step, by selecting a tag from a given resource-specific tag cloud, the system again presents a paginated list of resources and the user might continue the navigation process in the same manner as before.

7.4 Problems of Tag Cloud Navigation

Resource-specific tag clouds are a simple way to connect resources within a tagging system, i.e. in a typical tagging system one can find nearly 99% of the resources interlinked with each other within a tag cloud network [11]. However, this simple approach to building tag clouds exhibits certain problems. In particular, resource-specific tag clouds are vulnerable to a so-called pagination effect [11]. In other words, by k-limiting the resource list of a given tag (with typical pagination values such as 5, 10, or 20), the connectivity of the tag cloud network collapses drastically. Practically, this leads to a situation where the tag cloud network consists of isolated network clusters (components) that are not linked to each other anymore. In other words, the users cannot reach one network fragment from another network fragment by navigating resource-specific tag clouds. One simple solution to this problem is to select resource for inclusion in a k-limited resource list uniformly at random [11]. For example, whenever the user clicks on a given tag in the tag cloud the system randomly selects k resources and presents them to the user. This leads to situation, that not always the same links are selected which leads to the situation that isolated network clusters are created [11]. As [4, 11] have shown this approach produces a random network that is, even for small values of k, completely connected.

7.4.1 Navigable vs. Efficiently Navigable Tag Cloud Networks

Another interesting issue in that context is the question if such randomly generated networks are also navigable. From a network-theoretic point of view Kleinberg [13, 14, 15] showed that a navigable network can be formally defined as a network with a low diameter [21] bounded by $log(N)$, where N are the number of nodes in the network, and an existing giant component, i.e. a strongly connected component containing almost all nodes. Additionally, Kleinberg defined an "efficiently" navigable network as a network possessing certain structural properties so that it is possible to design efficient decentralized search algorithms (algorithms that only have local knowledge of the network) [13, 14, 15]. The delivery time (the expected number of steps to reach an arbitrary target node) of such algorithms is polylogarithmic or at most sub-linear in N. Put short, in [15] Kleinberg also showed that naive random networks algorithms form network structures which require linear search time ($O(N)$), i.e. in the worst case one has to visit all N nodes within a network to reach a certain destination node, i.e. such networks are not efficient navigable. However, in [15] Kleinberg also showed that hierarchical network models generate networks which are navigable in polynomial of $O(logN)$. Thus, we applied a hierarchical network model to generate tag clouds in the Austria-Forum with the goal to support efficient navigation within the system.

7.5 Algorithm

7.5.1 Hierarchical Tag Cloud Construction Algorithm

Hierarchical network models [15] are based on the idea that, in many settings, the nodes in a network can be organized in a hierarchy. The hierarchy can be represented as a b-ary tree and network nodes can be attached to the leaves of the tree. For each node v, we can create a link to all other nodes w with the probability p that decreases with $h(v, w)$ where h is the height of the least common ancestor of v and w in the tree. Networks generated by this model are "efficiently" navigable [15].

To some extent, we showed that a hierarchical network model can be applied for the creation of efficiently navigable tagging systems [11]. For that purpose we developed a hierarchical network generator that 1) sorts the resource list of a given paginated tag by frequency, 2) creates resource clusters of size 10 by traversing the sorted resource list sequentially, 3) creates a balanced b-ary (b = 5) tree where the number of leaves is equal to the number of the resource clusters, 4) traverses the tree in postorder from left to right and attaches resource clusters to the tree leaves, and 5) uses this tree structure to obtain the link probability distribution for connecting a resource-

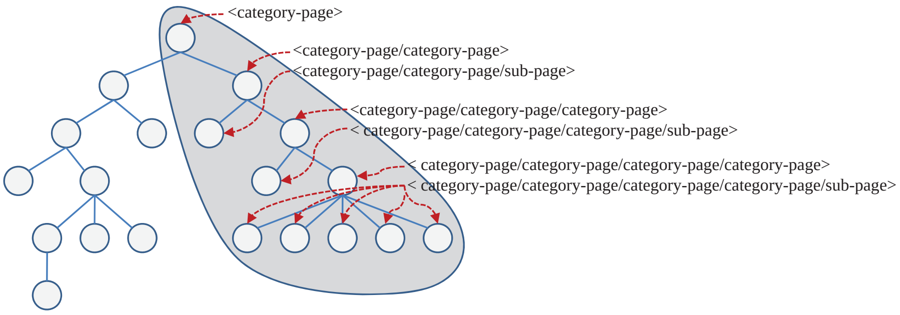


Figure 7.1: Hierarchical structure and URL addressing schema within the Austria-Forum.

Algorithm 3 Hierarchical Tag Cloud Construction Algorithm

getTagCloud: url, n
if (url is category-page) **then**
  $TC_r^n \leftarrow$ select top n tags sorted by t_f where r.url.startsWith(url)
else
  $TC_r^n \leftarrow$ select top n tags sorted by t_f
end if
return TC_r^n

specific tag node with resources of a given paginated tag. However, the main issue of this simple idea is, that the the tree creation process follows the statistical properties of the tagging dataset only, i.e. the generated hierarchy does not follow any semantic relations.

A simple idea to retain the semantics of the system is to reuse the given hierarchical data organization schema of the system. Since, the Austria-Forum organizes pages within the system into categories, sub-categories, sub-sub-categories, etc. and pages (see Figure 7.1), we designed in previous work an algorithm that generates tag clouds in a recursive and hierarchical manner (see Algorithm 3) based on the given resource structure of the Austria-Forum. Without evaluating the algorithm against the pagination effect, we argued for the efficiency of this approach due to it's nature of linking documents in a hierarchical manner.

7.5.2 Hierarchical Resource List Generation Algorithm

A possible better idea to address the pagination effect in tagging systems and also to apply Kleinberg's hierarchical network model is to reuse the given hierarchical organization schema of the system as the basis for generating a link probability distribution p to construct the resource lists of the tagging system (see also [11]). Since the hierarchical network model as introduced

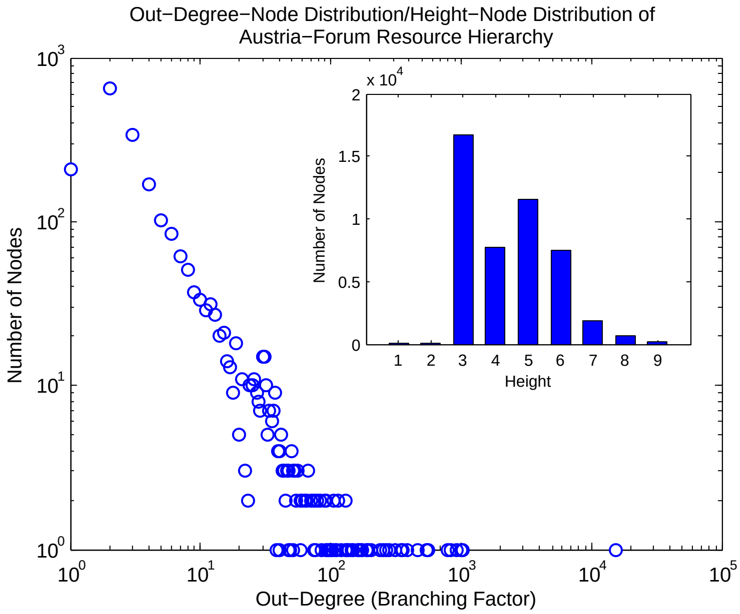


Figure 7.2: Out-degree distribution and node distribution of the Austria-Forum resource hierarchy.

by Kleinberg assumes a complete and balanced tree of the resources whereas typical hierarchically structured Web content does not follow this premiss (see Figure 7.2) an algorithm implementing Kleinberg's approach has to work with approximations.

The intuition which we followed with our algorithm is that the probability that an article is linked with other articles from the same category is higher than the probability that an article is linked with articles from other categories (cf. [33, 1, 15]). Put short, this can be modeled by defining a link selection function that inter-links two nodes (articles) v, w according to a link probability function that is equal to $p = e^{-dist(v,w)}$ (cf. [33]) and a distance function that is calculated as $dist(v, w) = h_v + h_w - 2h(v, w) - 1$, where h_v, h_w are the heights of two nodes v, w in the hierarchy and where $h(v, w)$ is the height of the least common ancestor of the nodes v, w in the hierarchy (cf. [1]). In Algorithm 4 the actual algorithm is presented.

7.6 Evaluation

To evaluate the presented algorithms, we conducted two types of experiments. The first experiment is based on a network-theoretic simulation approach that integrates the following two modules:

Algorithm 4 Hierarchical Resource List Generation Algorithm

For any given node $r(t) \in R$ in the resource hierarchy R, where t is the tag applied to this node, we find all other nodes $r_j(t) \in R$ and calculate distance $dist(r(t), r_j(t)) = h(r(t)) + h(r_j(t)) - 2h(r(t), r_j(t)) - 1$. For all found nodes $r_j(t) \in R$ we put $r_j(t)$ according to the distance $dist(r(t), r_j(t))$ into clusters $cl_x = [r_i, ..., r_j]$ and store these clusters into an array $rdist(i)_{r(t)} = [cl_{dist_1}, ..., cl_{dist_{n-1}}]$. Now, to select k links from the resource list, we generate k random numbers $i_k = 1 \ldots \text{sizeof}(rdist(i)_{r(t)})$ with a probability density function $p = e^{-x}$ with $x = 1 \ldots \text{sizeof}(rdist(i)_{r(t)})$ and select k clusters $cl_{i_k} \in rdist(i_k)_{r(t)}$ returning for each cluster just one element which is selected uniform at random.

- a **network-theoretic module** based on the Stanford Snap[3] library to calculate and evaluate network properties such as the size of the Largest Strongly Connected Component (LSCC) or the Effective Diameter (ED) [11] of the tag cloud network
- and a **searcher module** which implements a hierarchical decentralized searcher to simulate tag cloud based navigation.

The second type of experiment is based on a controlled user study that evaluates our approach of hierarchically constructed tag clouds against a baseline.

7.6.1 Simulations

In the following sections we present the results of our simulations. Similar to our previous work [11], we model the tag cloud network of a tagging system as a bipartite hypergraph of the form $V = R \cup T$ [11], where R is the set of resources and T the set of tags. Since the resource lists are limited to a certain value k which forces the tag cloud network into a directed unipartite tag-resource network (with resource specific tags), we performed our evaluations onto the projected resource-resource network. In order to measure navigability, we calculated the size of the largest strongly connected component (LSCC) and the effective diameter (ED) on that network. As defined in Section 7.4.1, we consider navigable networks to be networks that have a low diameter bounded logarithmically and a giant component.

Datasets

To evaluate our algorithms overall 10 different types of tag cloud networks were generated. They all vary in the way how the tag clouds and the resource lists were calculated. Since one of our recent studies [11] showed that limiting

[3] http://snap.stanford.edu/

the tag cloud to practically feasible sizes (e.g. 5, 10, or more) does not influence navigability, we set the tag cloud size in our experiments to a fixed value of $n = 30$. For the purpose of evaluation, we varied the value k, i.e. the maximum number of links in the resource list, to $k = 15, 50$, which is expected to impair navigability [11].

- **Network NN (=Naive Naive)**: This tag cloud network is generated by the most commonly and naive tag cloud and resource list calculation approach used these days in tagging systems [11]. In other words, the tag cloud calculation algorithm in this setting follows a simple TopN approach displaying the most frequent n tags in the tag cloud while the resource list calculation algorithm sorts the resources descending chronological order and selecting the k most top resources.

- **Network NR (=Naive Random)**: This tag cloud network is generated by using a naive TopN algorithm (cf. Dataset G) for tag cloud calculations displaying the most frequent n tags in the tag clouds. The resource list is generated selecting k resources uniform at random.

- **Network NP (=Naive Popularity)**: This tag cloud network is generated by using the TopN tag cloud calculation algorithm. The resource list is calculated sorting the resources by popularity and selecting the k most top resources.

- **Network NS (=Naive Similarity)**: This tag cloud network is generated by using the TopN tag cloud calculation algorithm. The resource list is calculated sorting the resources by (cosine) similarity and selecting the k most top resources.

- **Network NH (=Naive Hierarchical)**: This tag cloud network is generated by using the TopN tag cloud and the hierarchical resource list generation algorithm introduced in Algorithm 4.

- **Network HN (=Hierarchical Naive)**: This tag cloud network is generated by using the hierarchical tag cloud calculation algorithm introduced in Algorithm 3. The resource list is calculated sorting the resources chronologically in descending order and selecting the k most top resources.

- **Network HR (=Hierarchical Random)**: This tag cloud network is generated by using the hierarchical tag cloud algorithm introduced in Algorithm 3. The resource list is calculated selecting k resources uniform at random.

- **Network HP (=Hierarchical Popularity)**: This tag cloud network is generated using the hierarchical tag cloud algorithm introduced in

Algorithm 5 Hierarchical Decentralized Searcher (cf. [1])

Searcher: resource-resource graph G, resource-hierarchy T, start node v, target node w
while v != w **do**
  $v_i \leftarrow$ get all adjacent nodes $\in G$ from v
  // finds closest node according to $dist = dist_{min}$
  // where $dist(v_i, w) = h(v_i) + h(w) - 2h(v_i, w) - 1$
  v $\leftarrow$ findClosestNode (v_i, T)
end while

Algorithm 3. The resource list is calculated sorting the resources by popularity and selecting the k most top resources.

- **Network HS (=Hierarchical Similarity)**: This tag cloud network is generated using the hierarchical tag cloud algorithm introduced in Algorithm 3. The resource list is calculated sorting the resources by (cosine) similarity and selecting the k most top resources.

- **Network HH (=Hierarchical Hierarchical)**: This tag cloud network is generated using the hierarchical tag cloud algorithm introduced in Algorithm 3 and the hierarchical resource list algorithm introduced in Algorithm 4.

Results

As shown in Table 7.1, the tag cloud networks generated by a none random resource list calculation approach, such as network NN, NP, NS, HN, HP and HS, are not navigable (taking user interface limitations into account). They do not show a giant component containing (nearly almost) all nodes of the network which makes these networks unnavigable from a network-theoretical perspective. Contrary to this, networks based on a random resource list generation approach, such as network NR, NH, HR and HH are navigable. Note, that networks based on a hierarchical tag cloud calculation algorithm are also not navigable, except they implement a random resource list generation approach. However, compared to all other networks these networks show an effective diameter which is significantly smaller.

Since the previous experiment only presented that networks NR, NH, HR and HH are navigable, but not, if they are also useful for efficient navigation, a searcher routine was developed to determine this network property again from a network-theoretic perspective. As shown by Kleinberg, an "efficiently" navigable network is a network possessing certain structural properties so that it is possible to design an decentralized search algorithm, i.e. an algorithms that only has local knowledge of the network, and whose

Name	TC-Algo.	R-Algo.	n	k	LSCC	ED	NAV
NN_15	TopN	Chron.	30	15	0.567002	5.99404	unnav.
NN_50	TopN	Chron.	30	50	0.761011	5.39847	unnav.
NR_15	TopN	Rand.	30	15	0.949983	5.93975	nav.
NR_50	TopN	Rand.	30	50	0.949983	5.03066	nav.
NP_15	TopN	Pop.	30	15	0.508194	6.45806	unnav.
NP_50	TopN	Pop.	30	50	0.724650	4.84473	unnav.
NS_15	TopN	Sim.	30	15	0.597815	6.99932	unnav.
NS_50	TopN	Sim.	30	50	0.814015	5.8751	unnav.
NH_15	TopN	Hier.	30	15	0.949983	5.90015	nav.
NH_50	TopN	Hier.	30	50	0.949983	5.1389	nav.
HN_15	TopN-H	Chron.	30	15	0.566008	3.47673	unnav.
HN_50	TopN-H	Chron.	30	50	0.755314	2.93258	unnav.
HR_15	TopN-H	Rand.	30	15	0.968034	3.73302	nav.
HR_50	TopN-H	Rand.	30	50	0.968034	3.17498	nav.
HP_15	TopN-H	Pop.	30	15	0.462471	2.89974	unnav.
HP_50	TopN-H	Pop.	30	50	0.670375	2.89933	unnav.
HS_15	TopN-H	Sim.	30	15	0.640873	5.11472	unnav.
HS_50	TopN-H	Sim.	30	50	0.850880	3.90542	unnav.
HH_15	TopN-H	Hier.	30	15	0.968034	3.46743	nav.
HH_50	TopN-H	Hier.	30	50	0.968034	2.92611	nav.

TC-Algo. = Tag Cloud Calculation Algorithm, R-Algo. = Resource List Calculation Algorithm, Chron. = Chronologically Sorted, Rand. = Randomly Sorted, Pop. = Sorted by Popularity, Sim. = Sorted by Similarity, Hier. = Hierarchically Sorted, LSCC = Largest Strongly Connected Component, ED = Effective Diameter, NAV = Navigability, TopN-H = TopN Hierarchically Calculated, unnav. = unnavigable, nav. = navigable

Table 7.1: Tag cloud network dataset statistics: Largest Strongly Connected Component, Efficient Diameter and Navigability.

delivery time (the expected number of steps to reach an arbitrary target node) is poly-logarithmic or at most sub-linear in N, where N are the number of nodes in the network [14, 13, 15]. To that end, we implemented a hierarchical-decentralized searcher based on the ideas of [1] to evaluate the actual efficiency of our networks [10]. As input for the searcher the given resource hierarchy from the Austria-Forum was used. In Algorithm 5, the searcher as it was developed is presented. In words, the algorithms works as follows:

To find a certain target resource w from a certain start node v within the network, the searcher first selects all adjacent nodes v_i for the start node and then selects the node v from the network that has the shortest distance

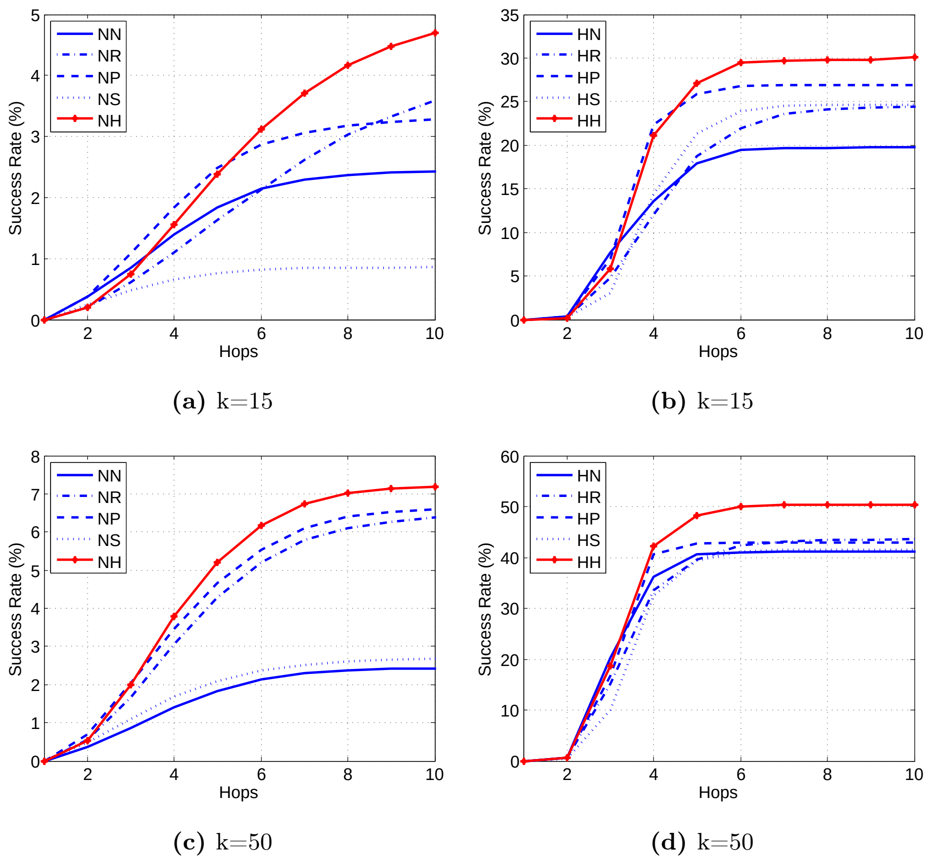


Figure 7.3: Hierarchical Decentralized Searcher cumulative hop-distributions for different values of k (size of the resource list).

$dist(v_i, w) = h(v_i) + h(w) - 2h(v_i, w) - 1$ to w node in the resource taxonomy T, with $h(v_i), h(w)$ being the heights of the two nodes v_i, w in the hierarchy and with $h(v_i, w)$ being the height of the least common ancestor of the two nodes v_i, w in the hierarchy [1]. In the next step, the adjacent nodes of v are again selected and the distances $dist(v_i, w)$ are calculated, while the node v with shortest distance is selected in the end. The process is continued until the target node w is reached.

In order to get statistically significant results, we simulated 100,000 search-requests starting randomly selected at a certain resource v_i and targeting at certain randomly selected resource w_i in the tag cloud network. Note, only search pairs v_i, w_i were considered for the simulations for which a path (v_i, w_i) was present in the network. The upper limit for a search was set to a value of maximum 10 hops in the simulations, i.e. we canceled searches which took more than 10 hops to find a target node w_i. If the searcher was not able to find a path further in the tag cloud network, we canceled the search task as well. If a search task was being canceled, we did not reset the searcher to find a new path for the same search pair v_i, w_i.

Even our previous experiments showed that the networks NN, NP, NS, HN, HP and HS are not navigable, we performed our simulations also on these tag cloud networks. This was done to gain more insights on the performance of such networks. To get comparable results, we extracted the LSCC of all networks and performed our simulations on the resulting networks.

As shown in Figure 7.3, tag cloud networks generated by a naive TopN tag cloud calculation algorithm produce poor results for a hierarchical decentralized search routine in such networks. The success rate is in the best case 4.5%-7% for $k = 15$ and 50. On the other hand, tagging systems implementing a hierarchical tag cloud algorithm perform much better in finding paths from a given start resource v_i to a certain target resource w_i (see right figures in Figure 7.3). For $k = 15$ we can reach in the best case 30% of all resources and for $k = 50$ over 50%. Interestingly, this is only the case, if we use a hierarchical resource list generation approach.

Hence, putting the results of the two experiments together, we can be seen that the best results can be obtained by combining hierarchical tag cloud construction with hierarchically constructed resource lists.

7.6.2 User Study

To quantify the usefulness of our approach not only with simulations but also empirically, we conducted additionally a small user study to confirm our findings.

Preliminaries

To prepare for the experiment, the latest (October 16, 2010) tagging dataset was downloaded from the Austria-Forum live system. Since user studies are typically time intense, we decided to compare our hierarchical tag cloud construction approach not against all other approaches as discussed before, but against the most popular approach for constructing tag clouds in tagging systems. To that end two different types of tag cloud networks were generated:

- **Network N:** Uses the TopN tag cloud calculation algorithm and the reverse chronologically sorted resource list calculation algorithm.
- **Network H:** Used the hierarchical tag cloud algorithm introduced in Algorithm 3 and the hierarchical resource list algorithm introduced in Algorithm 4.

The max. tag cloud size used for the experiment was $n = 30$ tags. The pagination factor used for generating the resource lists was $k = 50$.

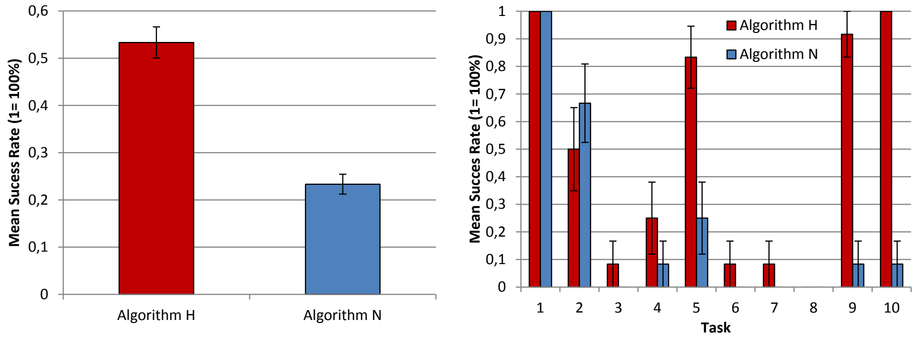
After that, we 10 selected uniform at random resource pairs (start and target resources) present in both tag cloud networks. The resource pairs were selected such that the targets were reachable in a minimum number of one, two, three, four and five steps (one step = two clicks, one for opening the resource list and for moving on to the next resource) and in *both* networks. The maximum of five steps was chosen since it was calculated that almost all resources were reachable in a minimum of five steps in both networks. The different step (path) lengths were selected to ensure that participants would have to navigate via a certain number of intermediate resources to reach their designated target resource. Finally, ten online tasks, one for each resource pair, were designed and directly implemented into the Austria-Forum system.

Procedure

In order to measure the performance of the both algorithms, a between-groups (independent measures) design was used for the experiment. All 24 participants were given the exact same 10 tasks (one for each pre-calculated resource pair). Users were asked to surf with different tag networks generated by different resource list generation algorithm. 12 users had to navigate in the pre-calculated tag Network H and 12 users had to navigate in the pre-calculated tag Network N. For each task (= one resource pair, start and target resource), the users were asked to reach the given target resource as fast as possible, using exclusively tags and the corresponding resource lists for navigation. For each task, participants were given a maximum of three minutes to reach the given target resource (this was found to be an appropriate upper limit during the pilot test phase). If the user could not find the target after the time has elapsed, the user was asked to cancel the search and continue with the next task.

Participants

All in all, 24 participants were invited to join the experiment, 16 of them male and 8 of them female. The median age of the users was 33 years, ranging from 22 to 56. All participants were experienced computer (on average 46 hours per week) and Internet users (on average 19 hours per week). 12 of them were experienced with the test system. Hence, in order to get valid results, they were split up into two groups, i.e. six of them were assigned to evaluate the chronologically sorting resource list generation approach and six of them were assigned to evaluate the hierarchical resource list generation algorithm. The study was performed at Graz University of Technology, Austria from November 8 to 12, 2010.



(a) Mean success rate overall tasks. **(b)** Mean success rate per task.

Figure 7.4: Mean success rates.

Results

To compare the two tag cloud construction algorithms with each other the success rate was measured. The number of clicks or time to reach the target resources was not evaluated since it was observed the success rates for navigating Network N were very low compared to Network H.

In Figure 7.4 (a), the average success rate for the both tag cloud networks H and C are presented. As shown, the mean user's success rate over all tasks with Network H is significantly higher ($p < 0.05$) than with tag Network N. More precisely, the experiment showed that the mean success rate for Network N is 23.3% while the mean success rate overall users for Network H is 53.3%. In other words, a user was more than twice as likely to be able to find a target in the tagging system that has implemented a hierarchical resource list generation algorithm than in the tagging system that has been implemented a chronologically sorting resource list generation algorithm. These results confirm the erlier results from Section 7.6. For exact the same tag cloud networks and resource pairs, simulations show an overall success rate for Network N of 20% and for Network H of 50%.

Figure 7.4 (b) shows the mean success rate for each individual task. For all tasks with path length > 1, tag Network H performed on average better than tag Network N. Significant differences were found for Tasks 5, 9, and 10. It was observed that the low success rates for Tasks 3, 6, 7, and 8 were due the high branching factors in the resource taxonomy for the target resources used for these tasks. In latest research [29], we have focused on that issue and found a simple and promising way to automatically construct resource taxonomies from tagging data with fixed branching factors. First simulations show that the overall success rate to navigate the tag network could be increased to 98%.

7.7 Related Work

In related research on tagging systems, tag clouds have been characterized as a way to translate the emergent vocabulary of a folksonomy into social navigation tools [26, 6]. Social navigation itself represents a multi-dimensional concept, covering a range of different issues and ideas. A distinction between direct and indirect social navigation, for example, highlights whether navigational clues are provided by direct communication among users (e.g. via chat), or whether navigational clues are indirectly inferred from historical traces left by others [20]. Based on this distinction, our work only focuses on indirect social navigation in the sense that it studies the effectiveness of traces ("tags") left by users in tagging systems. Other types of social navigation emphasize the need to show the presence of others users, to build trust among groups of users, or to encourage certain behavior [20].

Researchers have discussed the advantages and drawbacks of tag clouds, suggesting that tag clouds are a useful mechanism when users' search tasks are general and explorative (for example, learn about Web 2.0), while tag clouds provide little value for specific information-seeking tasks (for example, navigate to www.cnn.com) [26]. While the paper at hand focuses on network-theoretic aspects, cognitive aspects of navigation have been studied previously using, for example, SNIF-ACT [7] and social information foraging theory [23]. Other work has studied the motivations of users for tagging [16], and how they influence emergent semantic (as opposed to navigational) structures. The navigational utility of single tags has been investigated [5] with somewhat disappointing results. With time the tags become harder and harder to use as they lose specificity and reference too many resources. Such tags are exactly those paginated tags where new pagination algorithms are needed.

Navigation models for tagging systems have been also discussed recently. In [24] authors describe a navigation framework for tagging systems. The authors apply the framework to analyze possible attacks on tagging systems. In principle, the framework identifies a navigation channels as any combination of the basic elements of a tagging system (users, tags, and resources). Thus, the specific combination which we investigated in this paper can be summarized as the resource-tag or tag-resource navigation channel.

Recent literature also discusses further algorithms for the construction of tag clouds. The ELSABer algorithm [18] represents an example of such an effort aimed towards identifying hierarchical relationships between annotations to facilitate browsing. The work by [3] is another example, introducing entropy-based algorithms for the construction of interesting tag clouds. However, these algorithms have not found wide-spread adoption in current social tagging systems, and their usefulness to support navigation is largely unknown. In future work, it would be interesting to compare additional tag cloud construction algorithms with our approach. In addition, empirical studies of tagging systems have for example focused on comparing navigational characteristics of tag distributions to similar distributions produced by library terms [12].

7.8 Conclusions and Future Work

The main contribution of this paper is the introduction of a novel approach for interlinking resources in hierarchically-structured Web content. Based on a review of tag cloud limitations and an existing hierarchical algorithm for the construction of efficiently navigable networks, we discussed, implemented, and evaluated by simulations and a small user study a new approach to tag cloud construction. As shown the proposed algorithm creates tag-network structures which are more navigable than current state-of-the-art approaches for tag cloud and resource list construction.

References

[1] L. Adamic and E. Adar. How to search a social network. *Social Networks*, 27(3):187 – 203, 2005.

[2] M. Ames and M. Naaman. Why we tag: Motivations for annotation in mobile and online media. In *Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI 2007)*, CHI'07, pages 971–980, USA, NY, 2007. ACM.

[3] K. Aouiche, D. Lemire, and R. Godin. Web 2.0 OLAP: From Data Cubes to Tag Clouds. *4th International Conference, WEBIST 2008*, 18, 2008.

[4] B. Bollobás and F. R. K. Chung. The diameter of a cycle plus a random matching. *SIAM J. Discret. Math.*, 1(3):328–333, 1988.

[5] E. H. Chi and T. Mytkowicz. Understanding the efficiency of social tagging systems using information theory. In *HT '08: Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*, pages 81–88, New York, NY, USA, 2008. ACM.

[6] A. Dieberger. Supporting social navigation on the world wide web. *Int. J. Hum.-Comput. Stud.*, 46(6):805–825, 1997.

[7] W.-T. Fu and P. Pirolli. Snif-act: a cognitive model of user navigation on the world wide web. *Hum.-Comput. Interact.*, 22(4):355–412, 2007.

[8] T. Hammond, T. Hannay, B. Lund, and J. Scott. Social bookmarking tools (i): A general review. *D-Lib Magazine*, 11(4), 2005.

[9] Y. Hassan-Montero and V. Herrero-Solana. Improving tag-clouds as visual information retrieval interfaces. In *InScit2006: International Conference on Multidisciplinary Information Sciences and Technologies*, 2006.

[10] D. Helic and M. Strohmaier. Building directories for social tagging systems. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, CIKM '11, pages 525–534, New York, NY, USA, 2011. ACM.

[11] D. Helic, C. Trattner, M. Strohmaier, and K. Andrews. On the navigability of social tagging systems. In *Proc. of 2010 IEEE International Conference on Social Computing*, pages 161–168, Los Alamitos, CA, USA, 2010. IEEE Computer Society.

[12] P. Heymann, A. Paepcke, and H. Garcia-Molina. Tagging human knowledge. In *Proceedings of the third ACM international conference on Web search and data mining*, WSDM '10, pages 51–60, New York, NY, USA, 2010. ACM.

[13] J. Kleinberg. The small-world phenomenon: An algorithmic perspective. In *Proceedings of the 32nd ACM Symposium on Theory of Computing*, 2000.

[14] J. M. Kleinberg. Navigation in a small world. *Nature*, 406(6798), August 2000.

[15] J. M. Kleinberg. Small-world phenomena and the dynamics of information. In *Advances in Neural Information Processing Systems (NIPS) 14*, page 2001. MIT Press, 2001.

[16] C. Körner, D. Benz, A. Hotho, M. Strohmaier, and G. Stumme. Stop thinking, start tagging: Tag semantics arise from collaborative verbosity. In *19th International World Wide Web Conference (WWW2010), Raleigh, NC, USA, April 26-30, ACM*, 2010.

[17] C. Körner, R. Kern, H. P. Grahsl, and M. Strohmaier. Of categorizers and describers: An evaluation of quantitative measures for tagging motivation. In *21st ACM SIGWEB Conference on Hypertext and Hypermedia (HT 2010), Toronto, Canada, ACM*, Toronto, Canada, June 2010.

[18] R. Li, S. Bao, Y. Yu, B. Fei, and Z. Su. Towards effective browsing of large scale social annotations. *Proceedings of the 16th international conference on World Wide Web*, page 952, 2007.

[19] C. Marlow, M. Naaman, D. Boyd, and M. Davis. Ht06, tagging paper, taxonomy, flickr, academic article, to read. In *Proc. Seventeenth Conference on Hypertext and Hypermedia (Hypertext 2006)*, HT'06, pages 31–40, USA, NY, 2006. ACM.

[20] D. Millen and J. Feinberg. Using social tagging to improve social navigation. In *Workshop on the Social Navigation and Community Based Adaptation Technologies. Dublin, Ireland.* Citeseer, 2006.

[21] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[22] O. Nov and C. Ye. Why do people tag?: motivations for photo tagging. *Commun. ACM*, 53:128–131, July 2010.

[23] P. Pirolli. An elementary social information foraging model. In *Proceedings of the 27th international conference on Human factors in computing systems*, pages 605–614. ACM, 2009.

[24] M. Ramezani, J. Sandvig, T. Schimoler, J. Gemmell, B. Mobasher, and R. Burke. Evaluating the impact of attacks in collaborative tagging environments. In *Computational Science and Engineering, 2009. CSE '09. International Conference on*, volume 4, pages 136 –143, aug. 2009.

[25] A. Rivadeneira, D. Gruen, M. Muller, and D. Millen. Getting our head in the clouds: toward evaluation studies of tagclouds. *Proceedings of the SIGCHI conference on Human factors in computing systems*, page 998, 2007.

[26] J. Sinclair and M. Cardew-Hall. The folksonomy tag cloud: when is it useful? *Journal of Information Science*, 34:15, 2008.

[27] M. Strohmaier, C. Koerner, and R. Kern. Why do users tag? Detecting users' motivation for tagging in social tagging systems. In *International AAAI Conference on Weblogs and Social Media (ICWSM2010), Washington, DC, USA, May 23-26*, 2010.

[28] C. Trattner. Improving tag cloud navigability: A comparative study. In *Proc. of the 33th International Conference on Information Technology Interfaces*, Cavtat, Dubrovnik, Croatia, 2011. IEEE.

[29] C. Trattner. Improving the navigability of tagging systems with hierarchically constructed resource lists and tag trails. *CIT*, 19(3):155–167, 2011.

[30] C. Trattner, H. Denis, and M. Strohmaier. On the construction of efficiently navigable tag clouds using knowledge from structured web content. *Journal of Universal Computer Science*, 2011.

[31] C. Trattner, I. Hasani-Mavriqi, D. Helic, and H. Leitner. The austrian way of wiki(pedia)!: development of a structured wiki-based encyclopedia within a local austrian context. In *Proceedings of the 6th International Symposium on Wikis and Open Collaboration*, WikiSym '10, pages 1–10, New York, NY, USA, 2010. ACM.

[32] C. Trattner and D. Helic. Extending the basic tagging model: Context aware tagging. In *IADIS International Conference WWWInternet 2009*, pages 76 – 83. IADIS, 2009.

[33] D. J. Watts, P. S. Dodds, and M. E. J. Newman. Identity and search in social networks. *Science*, 296:1302–1305, 2002.

[34] X. Wu, L. Zhang, and Y. Yu. Exploring social annotations for the semantic web. *Proceedings of the 15th international conference on World Wide Web*, page 426, 2006.

8 Enhancing the Navigability of Tagging systems with Tag Hierarchies

This chapter is based on the paper *"Enhancing the Navigability of Tagging systems with Tag Hierarchies"* which was presented at the 10th International Conference on Knowledge Management and Knowledge Technologies in 2011.

It continues the work on enhancing the navigability of tagging systems and reviews the potentials and limitations of tag hierarchies for efficient search and navigation in tagging systems. To that end, this chapter presents a novel algorithm for the creation of hierarchical structures that support efficient navigation in social tagging systems. We evaluate the proposed algorithm from a theoretical, semantic and empirical point of view. With these evaluations we are able to show a high performance and usefulness of the proposed approach.

The original contribution was published in the proceedings of the conference and can be found in [25].

8.1 Abstract

Tagging introduces an additional intuitive and easy method to organize resources in information systems. Although tags exhibit useful properties for e.g. personal organization of information, recent research has shown that the navigability of social tagging systems leaves much to be desired. When browsing social tagging systems users often have to navigate through huge lists of potential results before arriving at the desired resource. Thus, from a user point of view tagging systems are typically hard to navigate. To overcome this issue, we present in this paper a novel approach to support navigation in social tagging systems. We introduce tag-resource taxonomies that aim at to supporting efficient navigation of tagging systems. To that end, we introduce an algorithm for the generation of these hierarchical struc-

tures. We evaluate the proposed algorithm and hierarchies from a theoretical, semantic and empirical point of view. With these evaluations we are able to show a high performance and usefulness of the proposed idea.

8.2 Introduction

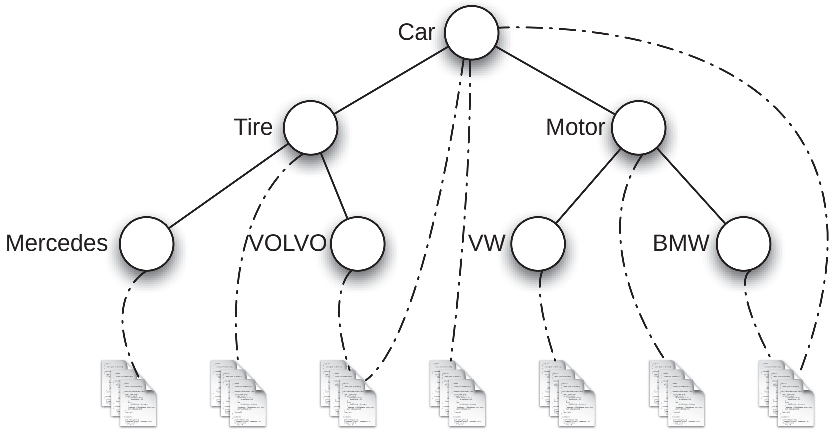
Tagging provides an easy and intuitive way to annotate, organize and retrieve resources on the web. For this reason, the popularity of social tagging systems has increased tremendously in recent years. To give some examples: Delicious[1] enables the annotation of personal bookmarks with tags, Flickr[2] allows users to describe their photos by tagging and Youtube[3] supports easier finding of videos via tags by content creators.

While there has been a lot of work on the structure of social tagging systems, little is known about the ways users use and navigate such systems. Some previous work by Chi et al. [4] observed that the navigability leaves much to be desired. There, the authors showed that the number of new tags does not grow as quickly as the number of tagged resources in mature social tagging systems such as BibSonomy, CiteULike or Delicious. Therefore a lot of tags exist that refer to a large number of documents within such systems. To illustrate this problem from a user perspective: when users click on a popular Delicious tag such as "`web`" they retrieve 6.5 million resources in reverse chronological order – thus, rendering the system unusable from a navigational point of view.

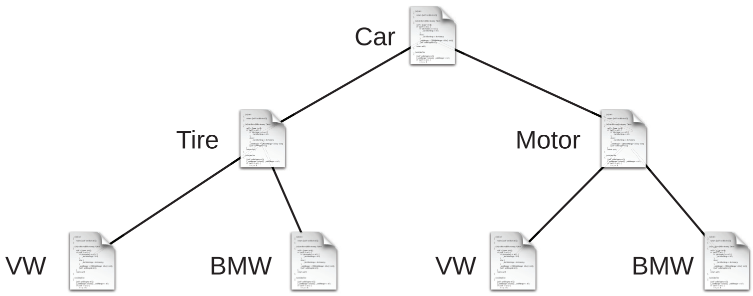
To overcome this issue recent research has investigated methods and strategies to make tagging systems more navigable. One prominent example of such endeavors are so-called tag taxonomies [14] – a method which allows the user to navigate to related concepts (tags) in a tagging system in a hierarchical and efficient manner (see also [12] for evaluation of several similar approaches). In this paper we introduce the notation of *tag-resource taxonomies*. Contrary to the idea of tag taxonomies, this approach enables the users not only to quickly navigate to related concepts but also to *resources* from a tagging system. With the approach of tag taxonomies, as it will be shown in this paper, efficient navigation to the resources of the tagging system is not possible. In a theoretical, semantic and empirical evaluation we show a high performance and usefulness of tag-resource taxonomies. To the best of our knowledge this is the first work that describes the notation of tag-resource taxonomies. Moreover, this approach significantly improves navigability of social tagging systems when compared with tag taxonomy approaches.

The paper is structured as follows: In Section 8.3 we introduce a novel

[1] http://www.delicious.com/
[2] http://www.flickr.com/
[3] http://www.youtube.com/



(a) Tag Taxonomy



(b) Tag-Resource Taxonomy

Figure 8.1: Tag Taxonomy vs. Tag-Resource Taxonomy.

approach to construct tag/res hierarchies and illustrate the algorithms that were created for this purpose. This is followed by Section 8.4 explaining our evaluation. Section 8.5 gives an overview of related work. Finally in Section 8.6 we conclude our findings and point to future work.

8.3 Approach

To tackle the issue of poor navigability in tagging systems, we introduce a novel approach to generate *tag-resource taxonomies*. The goal of the approach is to offer the user a simple tool to navigate the tagging system in an efficient way. According to Kleinberg [16], efficient navigation in a network is possible if all resources are navigable in a polynomial of $log(n)$, where n is the number of resources in a network. With the approach of tag-resource taxonomies, and as it is shown in Section 8.3.1, this prerequisite is fulfilled, i.e. it is possible to navigate a tagging system in a polynomial of $log(n)$.

Basically, a tag-resource taxonomy is a hierarchy containing both resources and tags. The basis of a tag-resource taxonomy is the so-called *resource taxonomy*. A resource taxonomy is a hierarchy where the resources of a tagging system are arranged in a unique and taxonomic way, i.e. each resource of the tagging system occurs only once and parent nodes are more general than their child nodes.

Given such a resource taxonomy we construct the final tag-resource taxonomy by using a labeling algorithm that applies tag information to each resource in a descriptive and general manner. Hence, each resource in the resource taxonomy has one tag label attached to describe the underlying resource. The resulting tag-resource taxonomy presented to the user is then a tag hierarchy where the tags refer to a constant number of resources.

Figure 8.1 gives an example of a tag taxonomy as compared to a tag-resource taxonomy. In a tag taxonomy tags appear only once in the hierarchy. However, resources can be referred by different tags. In a tag-resource taxonomy on the other hand resources occur only once while tags can appear on multiple and on different levels.

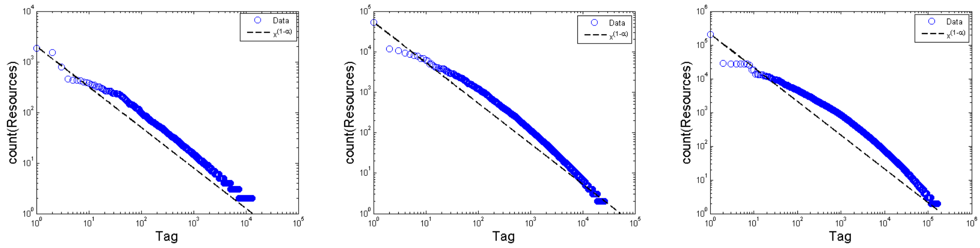
8.3.1 Why Usefulness of Tag Taxonomies for Navigation is Limited

A tag taxonomy allows the user to navigate to a designated tag (concept) efficiently, but navigation to a particular resource is still a problem due the so-called pagination effect. As shown by [10], in tagging systems the tag-resource distribution follows a power-law function (see Figure 8.2), i.e. there are many tags that refer to a large number of resources. In BibSonomy or CiteULike for instance there are tags, which refer to hundreds or even thousands of resources. To make such frequently used tags still usable for the user, developers typically paginate the result list of such tags by a certain factor k. Hence, in the worst case the user has to click through the whole paginated result list to find the desired resource. In detail, in the worst case the user would have to click

$$max\{click(T_{tag})\} = \frac{|max\{t\}|}{k} + max\{depth(T_{tag})\} \tag{8.1}$$

times to reach a designated target resource with the approach of a tag taxonomy.

The term $|max\{t\}|$ in Equation 8.1 describes the size of the most frequently used tag in the tagging system. The term k stands for the pagination factor and $max\{depth(T_{tag})\}$ denotes the maximum depth of the tag taxonomy. As shown in [26], the size of the most frequently used tag can be estimated as $|max\{t\}| = c_1 \cdot |r|$, where c_1 is a constant typically ranging between $[0.1, ..., 0.2]$ and $|r|$ is the number of unique resources in the tagging system. $max\{depth(T_{tag})\}$ can be estimated as, $log_{b/2}\,|t|$, supposing that



(a) Austria-Forum **(b)** BibSonomy **(c)** CiteULike

Figure 8.2: Tag distributions.

T_{tag} is a complete and fixed branched tree with branching factor b. The factor $|t|$ describes the number of unique tags in the tagging system. $|t|$ can be estimated as $|t| = c_2 \cdot |r|$, where c_2 is a constant. Therefore, Equation 8.1 can be formalized as

$$max\{click(T_{tag})\} = \frac{c_1 \cdot |r|}{k} + log_{b/2}(c_2 \cdot |r|),\ b \geq 2 \tag{8.2}$$

or

$$max\{click(T_{tag})\} \approx \frac{c_1 \cdot |r|}{k} \tag{8.3}$$

supposing that $log_{b/2}(c_2 \cdot |r|) \ll \frac{c_1 \cdot |r|}{k}$.

By generating a tag-resource taxonomy the worst case scenario is significantly better, especially for large numbers of $|r|$. Suppose the tag-resource taxonomy T_{res} is complete and has a fixed branching factor b, with $b = k$. A user would have to click

$$max\{click(T_{res})\} = max\{depth(T_{res})\} = log_{k/2}\,|r|\,,\ k \geq 2 \tag{8.4}$$

times in the worst case to reach a designated target resource. Then for large values of $|r|$ we have:

$$log_{k/2}\,|r| \ll \frac{c_1 \cdot |r|}{k} \tag{8.5}$$

Hence, according to the definition of Kleinberg [16] (see Section 8.3), and contrary to tag taxonomies, tag-resource taxonomies allow the user to navigate to the resources of a tagging system in an efficient manner, i.e. in a polynomial of $log(n)$.

To give an example: Let us calculate the number of maximum clicks for the tag datasets shown in Table 8.1 and compare the resulting tag taxonomy and tag-resource taxonomy for $k = 10$. As shown in Table 8.2, in a tag taxonomy the user would have to click $max\{click(T_{tag})\} = 184$ times

Tag Dataset	Austria-Forum	BibSonomy	CiteULike
$\|r\|$	19,430	233,712	949,851
$\|t\|$	13,314	26,285	163,642
$\|max\{t\}\|$	1,838	52,777	207,990
α	2.2	1.9	2.0

Table 8.1: Statistics of Austria-Forum, BibSonomy and CiteULike tag dataset.

	Austria-Forum	BibSonomy	CiteULike
$max\{click(T_{tag})\}$	184	5,278	20,799
$max\{click(T_{res})\}$	6.1	7.7	8.5

Table 8.2: Tag Taxonomy vs. Tag-Resource Taxonomy: Maximum number of clicks.

in the Austria-Forum tag dataset, respectively $5,278$ and $20,799$ clicks in the BibSonomy and CiteUlike tag dataset, to reach a desired resource in the worst case. In a tag-resource taxonomy the worst case would only be $max\{click(T_{res})\} = 6.1$ clicks for the Austria-Forum dataset or 7.7 clicks for the BibSonomy and 8.5 clicks for the CiteULike tag dataset.

Now, in order to calculate the number of tags suffering from the pagination effect we can define the following equations: Since we know that the tag distribution (see Figure 8.2) has power-law qualities we can approximate the number of paginated tags $|t_p|$ as follows [5]

$$r_i = \frac{\alpha - 1}{t_{min}} \cdot \left(\frac{t_i}{t_{min}}\right)^{-\alpha}, \ t_{min} > 0 \tag{8.6}$$

The parameter α can be approximated with the method of maximum likelihood as

$$\alpha \cong 1 + |t| \left[\sum_{i=1}^{|t|} ln \frac{t_i}{t_{min}}\right]^{-1} \tag{8.7}$$

With $r_i = k$ and $t_{min} = 1$, resolved by t_p the Equation 6 can be re-written as

$$t_p = \left(\frac{\alpha}{k} - \frac{1}{k}\right)^{\left(\frac{1}{\alpha}\right)} \tag{8.8}$$

The number of paginated tags $|t_p|$ can be then calculated as

$$|t_p| = |t| \cdot \left(\frac{\alpha}{k} - \frac{1}{k}\right)^{\left(\frac{1}{\alpha}\right)} \tag{8.9}$$

Example: Let us calculate the number of paginated tags for the tag datasets as shown in Table 8.1 for $k = 10$. Then, as shown in Table 8.3,

	Austria-Forum	BibSonomy	CiteULike
$\lvert t_p \rvert$ (%)	5079 (38%)	7401 (28%)	51748 (32%)

Table 8.3: Number of paginated tags.

Name	b	n	$max\{click(T_{res})\}$	$mean\{click(T_{res})\}$
Res2	2	19,430	17	12.45
Res5	5	19,430	10	5.93
Res10	10	19,430	8	4.44

Table 8.4: $max\{click(T_{res})\}$ and $mean\{click(T_{res})\}$ for different branching factors b.

within the Austria-Forum dataset 38% of all tags suffer from the pagination effect, respectively 28% in the BibSonomy tag dataset and 32% in the CiteULike tag dataset. Or in other words, for a commonly used resource list of the length of $k = 10$, nearly 1/3 of all tags suffer from the so-called pagination effect, i.e. the resources of such tags are not navigable in an efficient way!

8.3.2 Description of the Algorithm

Resource Taxonomy Generation Algorithm

As described in Section 8.3, the basis of the tag-resource taxonomy is the so-called resource taxonomy – a taxonomy where the resources of the tagging system are arranged in a taxonomic manner. In order to generate a resource taxonomy from tagging data we developed Algorithm 6. In words, the algorithm works as follows:

The algorithm takes a tag dataset and the desired taxonomy branching factor as input parameters. Since the algorithm should generate a resource taxonomy with the most general resource of the tagging system as root node and related and less general resources as children, the algorithm calculates in the first step degree centrality for all resource of the supplied tagging dataset and stores the centrality-resource pairs into a map C. Degree centrality was chosen since, on the one hand, it is computed fast, and on the other hand, it was observed in our previous research [3] that degree centrality in tagging systems is highly correlated to sophisticated centrality measures such as closeness or betweenness centrality. In the next step, the algorithm sorts the resources in C according to their centrality values in descending order.

Subsequently, the algorithm takes the first element of C (i.e. the most general resource) and sets that resource as the root node of the resource taxonomy. Thereafter, the algorithm starts iterating trough the elements (resources) already present in resource taxonomy. For each resource in the resource taxonomy the algorithm then calculates the most similar resources (see *getMoreLikeThis* in Algorithm 6). In our prototypical implementation of the algorithm we implemented this function as a method that calculates cosine similarity between all co-occurring resources taking also the $tf \cdot idf$ values of the tag concepts into account. Additionally, the function ensures that only resources are returned which are not already part of the constructed resource taxonomy. The results of this function are stored into a map SIM, with resources as key values and with the provided similarity values as corresponding map values. To account for resource generality we multiply resource similarity values with their corresponding centrality values. The final scores are normalized to fall into the range of [0...1]. After that, the resources in SIM are sorted by the scores in descending order. This procedure ensures that the resources in SIM are not only similar to the currently processed resource but also sorted by their generality values. Thereafter the algorithm appends a maximum of b resources to the currently processed resource. The algorithm stops, if no similar resources could be found anymore.

Note, due the fixed branching factor b the algorithm does not guarantee that all resources of the tagging dataset are contained in the resulting resource taxonomy. However, as it will be shown in Section 8.4, the probability that one or even more resources are missing is relatively small due to the high number of existing links between the resources of the resource-to-resource network of a given tag dataset. On the other hand, in a tag taxonomy the probability that one concept is missing is significantly higher. The reason for this behavior is the fact that the tag-to-tag network of a tagging system is typically substantially less connected.

Tag-Resource Taxonomy Generation Algorithm

To produce the final tag-resource taxonomy on the basis of a generated resource taxonomy we developed Algorithm 7. In general it is a labeling algorithm taking a given resource taxonomy and a tagging dataset as input parameters. Tag information is used as label data. The algorithm tries to apply labels to the given resource taxonomy in such a way, that they are uniquely distinguishable and the most descriptive ones for the given resource. The candidate tags are thereby ranked by the method of tag co-occurrence. However, since it can happen that resources in the resource taxonomy have the same tags in their parent tag trail, due to the lack of available tags in the tagging system, additional meta-data is taken into account. We used title information of the resources as an additional way for differentiation.

In words the algorithm works as follows: In the first step the algorithm calculates, for each resource in the resource taxonomy, a list of co-occurring tags of all resource tags and stores this list sorted in descending order into a map. After that, the algorithm enters a loop and traverses the resource taxonomy in left-order. In this loop the actual labeling procedure is performed. Basically, the labeling process looks as follows: For each resource in

Algorithm 6 Resource Taxonomy Generation Algorithm

INPUT: Tag Dataset D, Branching Factor b
OUTPUT: Resource Taxonomy T
$C \leftarrow$ new $HashMap[]$
$T \leftarrow$ new $Tree[]$
for each $r_i \in F$ **do**
    $C[r_i] \leftarrow$ calculate degree centrality
end for
$sortByValues(C)$
/*sort C by values in descending order*/
$T[0] \leftarrow C[0]$
$SIM \leftarrow$ new $HashMap$
for i = 0; i < sizeof(T); i++ **do**
    /*get all similar resources of $T[i]$ and store the resources as key values and the similarity values into SIM*/
    $SIM \leftarrow getMoreLikeThis(T[i])$
    for each $r_i \in SIM$ **do**
        $T[r_i] \leftarrow T[r_i] \cdot C[r_i]$
    end for
    /*sort the resources in SIM by values in descending order*/
    $sortByValues(SIM)$
    for j = 0; j < $sizeof(SIM)$ and $j < b$; j++ **do**
        $T[i].append(SIM[i])$
    end for
end for
return T

the resource taxonomy the corresponding co-occurrence vector is consulted and the first label, i.e. the most frequent tag, is tried to be applied to the currently processed resource. If the currently used candidate tag is already a part of the tag trail of the currently processed resource (see variable $trails$ in Algorithm 7) the next element, i.e. the next frequent tag label is chosen as candidate tag. If no uniquely distinguishable tag trail can be constructed, i.e. the candidate tag label from the co-occurrence vector is already present in the tag trail of the resource additional meta data is taken into account. In our case, title information of the currently processed resource is used. Note, since tag and title information could be identical the proposed method is not completely free of collisions. However, to fix this issue one could consider additional meta data information or other methods to generate a unique label such as appending an iterative number for each label that occurs more than once. The algorithm stops if all resources of the given resource taxonomy are labeled.

Figure 8.3 shows the branching factor distribution for a tag-resource taxonomy with branching $b = 5$ generated from the Austria-Forum tag dataset. For branching factor $b = 5$ the algorithm does not generate a complete $b-tree$ (from levels 1 to 4 the resulting tree is complete, for levels > 4 the

Algorithm 7 Tag-Resource Taxonomy Generation Algorithm

```
INPUT: Resource Taxonomy T, Tag Dataset D
OUTPUT: Tag-resource Taxonomy
COTAGS ← new HashMap[newArray[]]
for i = 0; i < sizeof(T); i++ do
  Ts ← getTags(T[i], D)
  for j = 0; j < sizeof(Ts); j++ do
    cotags ← getCoocTags(Ts[j], D)
    sort(cotags)
    remove all tags from cotags that are not contained in T[i]
    COTAGS[T[i]].add(cotags)
  end for
end for
trails ← new HashSet[]
for each r_i ∈ T do
  /*T is traversed in left-order*/
  pl ← getParentLabels(r_i)
  for each l_j ∈ COTAGS[r_i] do
    if !pl.contains(l_j) then
      if !(trails.contains(pl.toString() + l_j)) then
        T[r_i].applyLabel(pl)
        trails.add(pl.toString() + l_j)
      end if
    end if
    if T[r_i] has no label then
      T[r_i].applyLabel(getTitle(r_i))
    end if
  end for
end for
return T
```

tree is not complete). The reason for this behavior is the fact that in tag networks there are resources which are just connected to a few resources, i.e if the branching factor b is beneath this threshold the resulting taxonomy becomes incomplete.

8.4 Evaluation

Now, since we have shown in theory that the approach of the so-called tag-taxonomies allows the user to navigate to the resource of a tagging system in an efficient way, we will provide in the following section results of four different experiments to show the usefulness of the proposed approach also in a practical setting.

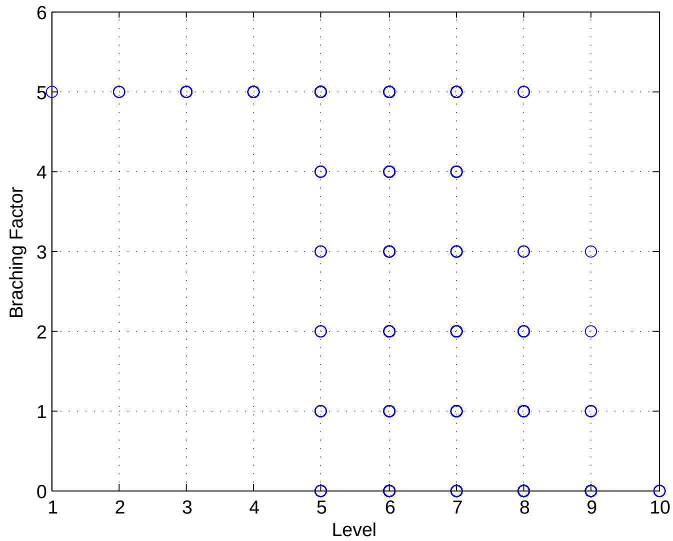


Figure 8.3: Example of a branching factor distribution for a tag-resource taxonomy with maximum branching $b = 5$.

8.4.1 Dataset

We used the tag dataset from a system called the *Austria-Forum*[4] [24] for the experiments. The Austria-Forum is a large online encyclopedia similar to Wikipedia providing the user with around $180,000$ resources on topics related to Austria. In contrast to Wikipedia, Austria-Forum offers an integrated tagging system, which allows users to assign tags to resources and to navigate to related resources via tag clouds. As of October 16^{th}, 2010, the Austria-Forum tag dataset contains $13,314$ tags, $19,430$ resources and $97,908$ tag assignments (see also Table 8.1).

8.4.2 Measuring the Average and Maximum Number of Clicks and the Drop Rate

In the first experiment we investigated

- the average and maximum tag-resource taxonomy depths for different branching factors b in order to measure the number of clicks a user would need to reach a designated target resource in the taxonomy and
- the number of missing resources (=drop rate) after the generation of a tag-resource taxonomy from tagging data with different branching factors b.

[4] http://www.austria-lexikon.at

Since the resulting tag-resource taxonomies are not complete, neither the average nor the maximum depth of the taxonomy can be estimated by formulas. If the tag-resource taxonomy was complete, we could calculate the maximum number of clicks as $max\{click(T_{res})\} = log_{b/2}(n)$, where n is the number of nodes in the taxonomy. Hence, these values were conducted empirically through an experiment.

For the experiment three different tag-resource taxonomies named $Res2$, $Res5$ and $Res10$ with three different branching factors $b = 2, 5$ and 10 were generated. In order to compare the resulting taxonomies against a golden standard taxonomy the DMOZ Open Directory Project (ODP) taxonomy[5] was consulted. This experiment was conducted to determine whether the generated tag-resource taxonomy would be usable or not.

As shown in Table 8.4 the tag-resource taxonomy with the smallest branching factor $b = 2$ is the deepest, needing a user $max\{click(T_{res})\} = 17$ clicks to reach a target resource in the worst case. On the other, and as expected the tag-resource taxonomy with highest branching factor $b = 10$ is less deepest taxonomy, i.e. in the worst case a user would have to click $max\{click(T_{res})\} = 8$ times to reach a desired resource. For $b = 5$ the $max\{click(T_{res})\} = 10$. On average for branching factor $b = 2$ the mean number of clicks is $mean\{click(T_{res})\} = 12.45$. For $b = 5$ the mean number of clicks is $mean\{click(T_{res})\} = 5.93$ and for $b = 10$ $mean\{click(T_{res})\} = 4.44$. The ODP Taxonomy has a mean depth of 6.86 [1]. The maximum depth is 13. Hence, compared with the ODP taxonomy the tag-resource taxonomy with branching factor $b \geq 5$ will be most usable.

In order to measure the number of missing resources (=drop rate) after the generation process of the taxonomies, we simply calculated the number of resources contained in tag-resource taxonomies $Res2$, $Res5$ and $Res10$ and compared it to the number of unique resources contained in the original Austria-Forum tag dataset. As shown in Table 8.4 and represented as parameter n, none of the resources dropped during the tag-resource taxonomy generation process. The reason for this behavior is the high number of existing links between the resources of the resource-to-resource network of the Austria-Forum tag dataset.

8.4.3 Measuring the Collision Rate

In the second experiment we measured the number of collisions when generating a tag-resource taxonomy with different branching factors b. As explained the tag-resource generation algorithm is not 100% collision free, i.e. it could happen that in a tag trail of a given resource the same tags occurs twice or even more often.

Hence, the goal of this experiment was to reveal how many collisions occur

[5] http://www.dmoz.org

Name	b	n	CR (%)
Res2	2	19,430	0.1%
Res5	5	19,430	0.2%
Res10	10	19,430	0.2%

Table 8.5: Collision Rates (CR) for different resource taxonomies with different branching factor b.

in general if a tag-resource taxonomy with a given branching b is created. For this experiment the three resource taxonomies from the former experiment were used. Table 8.5 shows the collision rates for the three generated tag-resource taxonomies. All in all, we observe that the collision rate is relatively small. However to make the approach totally free of collisions one might use additional meta-data as described in Section 8.3.2.

8.4.4 Measuring the Semantic Structure of the Tag-Resource Taxonomy

In the third experiment we measured the quality of the semantic structure of three tag-resource taxonomies that were generated for the two former experiments.

For that purpose, we consulted two different semantic measures – the Taxonomic F-Measure(in short TF) [6] and the Taxonomic Overlap(TO) [18]. Both measures identify the quality of a given taxonomy against a golden standard via common concepts. We used Germanet[6] as the golden standard for the experiment since the Austria-Forum tag dataset contains only German tags.

To determine the overall semantic quality of our three generated tag-taxonomies four tag taxonomies on the basis of the following popular tag taxonomy induction algorithms were generated – Hierarchical K-Means [7], Affinity Propagation [8, 21], Heymann [14] and Deg/Cooc [12, 2]. In the experiment, TF and TO values for all seven taxonomies were calculated and compared against each another. The goal of the experiment was to study how semantic structures generated by the tag-resource induction algorithm (Algorithm 2) compare to semantic structures produced by other popular tag taxonomy induction algorithms such as Hierarchical K-Means, Affinity Propagation, Heymann or Deg/Cooc [12]. Figure 8.4 shows the results of the semantic evaluation of the experiment. We observe that the higher the branching factor the better the semantic structure of the generated tag-resource taxonomies. The results indicate that tag-resource taxonomies with branching factors between $b = [5...10]$ perform on average as good as tag taxonomies based on a Affinity Propagation approach.

[6] http://www.sfs.uni-tuebingen.de/GermaNet/

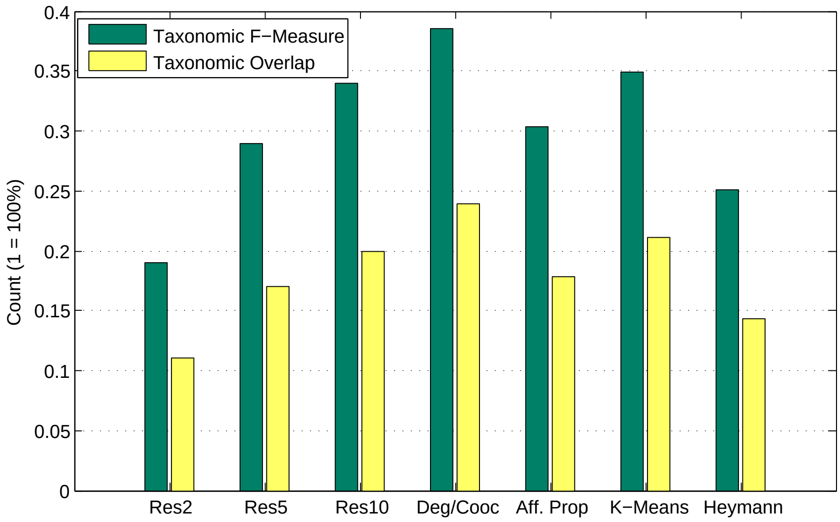


Figure 8.4: Results of the semantic evaluation of the three generated tag-resource taxonomies $Res2$, $Res5$ and $Res10$.

8.4.5 Empirical Analysis

In order to conduct whether or not the approach of a tag-resource taxonomy is also usable for humans, a user study based on the ideas of [23] was conducted. We used a tag-resource taxonomy with branching factor $b = 10$ for the following experiment.

First, we took the tag-resource taxonomy with branching factor $b = 10$ and extracted 100 tag trails uniformly at random from the tag-resource taxonomy. After that, a Deg/Cooc tag taxonomy with a maximum branching factor of $b = 10$ was generated in order to compare our approach of a tag-taxonomy to an existing method. Again, 100 tag trails were extracted uniformly at random from the generated tag taxonomy. Since shorter concept trails are typically better evaluated, we chose tag trails from both taxonomies that had a minimum tag trail length of 3 concepts (excluding the root node). After that, we presented the trails of both taxonomies in random order and generated an online test containing 200 tag trails, 837 relations and $1,037$ concepts. Each of our users were given exactly same tag trails. To insure that users knew how to evaluate the given tag trails a sample taxonomy with extracted tag trails was handed to them in addition to a detailed description of how to evaluate the trails. During the test the users were asked to rate the trails according to the classification schema presented in Table 8.6. All in all, 9 test subjects from three different departments at our university participated in the experiment. All participants were experienced computer users and familiar with user studies and the evaluation of concept hierarchies. The study was conducted online between April 25^{th} and 28^{th} of 2011.

Classification	Description
Correct	Correct hierarchy relation
Related	Correct relation, but not hierarchical or reverse hierarchical
Equivalent	Synonym
Not Related	The relations do not have anything to do with each other
Unknown	The evaluator does not recognize the meaning of the tag(s)

Table 8.6: Classification Labels for the User Evaluation.

Name	Corr. (%)	Rel. (%)	Equ. (%)	Not Rel. (%)	Unknown(%)
Deg/Cooc10	33.2	27.3	13	21.9	5.1
Res10	27.3	36.2	12.3	19.8	4.2

Table 8.7: Results of the empirical analysis of the tag-resource taxonomy with branching factor $b = 10$ compared to a Deg/Cooc tag taxonomy with branching factor $b = 10$ (Corr. = Correct, Rel. = Related, Equ. = Equivalent, Not Rel. = Not Related) .

Table 8.7 shows the results of the classification that was done by the study participants. These results indicate that the approach of a tag-taxonomy is not only useful in theory (as shown in Section 8.4.4) but also in practice. Compared to a tag taxonomy comprising only tags we can see that concept relations of a tag-resource taxonomy with branching factor $b = 10$ are only to 5% less hierarchically arranged than the tag concepts of the theoretically semantically most sound tag taxonomy induction approach the so-called Deg/Cooc tag taxonomy induction algorithm. Regarding the relatedness of the tag concepts we can observe that the tag-resource taxonomy was rated to 9% better than the Deg/Cooc tag taxonomy. Overall, the rating for the not related tags for both taxonomies was relatively small, taking into account that the maximum branching factor in both taxonomies was set to relatively small value of $b = 10$.

8.5 Related Work

For the presented work the following research topics on tagging are relevant:

8.5.1 Analysis of Social Tagging Systems

One of the first analysis of social tagging systems was conducted by Golder and Huberman [9]. In this work the authors show stable usage patterns within collaborative tagging systems and introduce an initial model of collaborative tagging. Subsequent work by Marlow et al. [19] introduces another model which gives insight into a simple taxonomy of incentives and contribution models within these systems. Hammond et al. [11] give a high level overview of different social tagging tools and examine various aspects such as audience and types of tagged media.

8.5.2 Navigation in Social Tagging Systems

As previously mentioned, Chi and Mytkowicz [4] studied Delicious using information theory (entropy) and found that the system becomes harder to navigate over time. The main reason for this is the small increase of tag vocabulary as opposed to the vast growth of tagging information in these systems. In previous work [13] we analyzed tag clouds as means of browsing tagging systems and showed that tag-resource networks have sufficient navigation properties in theory but also illustrated that user interface restrictions (such as pagination) spoil efficient navigation for all practical purposes.

8.5.3 Tag Semantics

In our own previous work [3] we compared different methods (such as network centrality, subsumption etc.) to measure the generality of tags in social systems. In [17] we showed that semantics within a social tagging system are heavily influenced by the users' tag usage. Users who are more verbose in the process of social tagging are better candidates for the construction of semantic structures out of folksonomies.

8.5.4 Creating Hierarchies from Social Tagging Data

Heymann et al. [14] converted a large corpus of tags annotating objects into a navigable hierarchical taxonomy of tags by evaluating the centrality of the tags in a similarity graph. In another work Solskinnsbakk et al. [23] constructed tag hierarchies using association rule mining of the corresponding tag set. Kiu and Tsui [15] introduced *TaxoFolk* - an algorithm which integrates tags and resources into a taxonomy by applying various data-mining techniques such as formal concept analysis. In another work Plangprasopchok et al. [20] propose a hierarchy generation algorithm based on an examination of user-defined relations within the system. Schmitz [22] gives insight into an algorithm that induces an ontology from tags in the Flickr system using a subsumption-based model. However, contrary to our work, none of these previous approaches examine the implications the resulting structures have on the navigability of the system.

8.6 Conclusions and Future Work

In this paper we introduced a novel approach to enhance the navigability of social tagging system through tag-resource taxonomies. We showed that tag taxonomies are in general well suited for finding related tag concepts, but perform worse in finding resources in an efficient number of clicks. By introducing the notation of the so-called tag-resource taxonomies we presented a method that tackles this issue. We illustrated in theory that with the approach of a tag-resource taxonomy it is possible to navigate to resources efficiently. Additionally to these findings, we evaluated the approach empirically and found that tag-resource taxonomies perform on a semantic level nearly as well or even better than other popular tag taxonomy approaches.

Thus, with the notation of tag-resource taxonomies we have introduced a novel hierarchical method that allows the user to navigate the resources in the tagging system in an efficient and semantically appropriate manner. To the best of our knowledge, this is the first work that describes such an efficient hierarchical navigation tool on the basis of tag-resource hierarchies.

References

[1] S. Alexaki, V. Christophides, G. Karvounarakis, D. Plexousakis, and K. Tolle. The ics-forth rdfsuite: Managing voluminous rdf description bases. In *SemWeb*, 2001.

[2] D. Benz, A. Hotho, and G. Stumme. Semantics made by you and me: Self-emerging ontologies can capture the diversity of shared knowledge. In *Proc. of the 2nd Web Science Conference (WebSci10)*, Raleigh, NC, USA, 2010.

[3] D. Benz, C. Körner, A. Hotho, G. Stumme, and M. Strohmaier. One tag to bind them all : Measuring term abstractness in social metadata. In *Proceedings of the 8th Extended Semantic Web Conference (ESWC 2011)*, Heraklion, Crete, May 2011.

[4] E. H. Chi and T. Mytkowicz. Understanding the efficiency of social tagging systems using information theory. In *HT '08: Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*, pages 81–88, New York, NY, USA, 2008. ACM.

[5] A. Clauset, C. R. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM Rev.*, 51:661–703, November 2009.

[6] K. Dellschaft and S. Staab. On how to perform a gold standard based evaluation of ontology learning. In *Proceedings of ISWC-2006 International Semantic Web Conference*, Athens, GA, USA, November 2006. Springer, LNCS.

[7] I. Dhillon, J. Fan, and Y. Guan. Efficient clustering of very large document collections. In R. Grossman, C. Kamath, and R. Naburu, editors, *Data Mining for Scientific and Engineering Applications*. Kluwer Academic Publishers, 2001.

[8] B. J. J. Frey and D. Dueck. Clustering by passing messages between data points. *Science*, 315(5814):972–976, January 2007.

[9] S. A. Golder and B. A. Huberman. The structure of collaborative tagging systems. *Journal of Information Science*, 32(2):198–208, 2006.

[10] H. Halpin, V. Robu, and H. Shepherd. The complex dynamics of collaborative tagging. In *Proceedings of the 16th international conference on World Wide Web*, WWW '07, pages 211–220, New York, NY, USA, 2007. ACM.

[11] T. Hammond, T. Hannay, B. Lund, and J. Scott. Social bookmarking tools (i): A general review. *D-Lib Magazine*, 11(4), 2005.

[12] D. Helic, M. Strohmaier, C. Trattner, M. Muhr, and K. Lermann. Pragmatic evaluation of folksonomies. In *Proc. of the 21st International World Wide Web conference, WWW '11*, New York, NY, USA, 2011. ACM.

[13] D. Helic, C. Trattner, M. Strohmaier, and K. Andrews. Are tag clouds useful for navigation? a network-theoretic analysis. *International Journal of Social Computing and Cyber-Physical Systems*, 1(1):33–55", 2011.

[14] P. Heymann and H. Garcia-Molina. Collaborative creation of communal hierarchical taxonomies in social tagging systems. Technical Report 2006-10, Stanford InfoLab, April 2006.

[15] C.-C. Kiu and E. Tsui. Taxofolk: A hybrid taxonomy-folksonomy structure for knowledge classification and navigation. *Expert Systems with Applications*, 38(5):6049 – 6058, 2011.

[16] J. M. Kleinberg. Small-world phenomena and the dynamics of information. In *Advances in Neural Information Processing Systems (NIPS) 14*, page 2001. MIT Press, 2001.

[17] C. Körner, D. Benz, A. Hotho, M. Strohmaier, and G. Stumme. Stop thinking, start tagging: Tag semantics arise from collaborative verbosity. In *19th International World Wide Web Conference (WWW2010), Raleigh, NC, USA, April 26-30, ACM*, 2010.

[18] A. Maedche and S. Staab. Measuring similarity between ontologies. In *Proc. Of the European Conference on Knowledge Acquisition and Management - EKAW-2002. Madrid, Spain, October 1-4, 2002*, volume 2473 of *LNCS/LNAI*, Heidelberg, 2002. Springer.

[19] C. Marlow, M. Naaman, D. Boyd, and M. Davis. Ht06, tagging paper, taxonomy, flickr, academic article, to read. In *Proc. Seventeenth Conference on Hypertext and Hypermedia (Hypertext 2006)*, HT'06, pages 31–40, USA, NY, 2006. ACM.

[20] A. Plangprasopchok and K. Lerman. Constructing folksonomies from user-specified relations on flickr. In *Proc. of 18th International World Wide Web Conference, WWW '09*, New York, NY, USA, May 2009. ACM.

[21] A. Plangprasopchok, K. Lerman, and L. Getoor. From saplings to a tree: Integrating structured metadata via relational affinity propagation. In *Proceedings of the AAAI workshop on Statistical Relational AI*, Menlo Park, CA, USA, July 2010. AAAI.

[22] P. Schmitz. Inducing ontology from flickr tags. In *Proceedings of the Workshop on Collaborative Tagging at WWW2006*, Edinburgh, Scotland, May 2006.

[23] G. Solskinnsbakk and J. Gulla. A hybrid approach to constructing tag hierarchies. In *On the Move to Meaningful Internet Systems, OTM 2010*, volume 6427 of *Lecture Notes in Computer Science*, pages 975–982. Springer Berlin / Heidelberg, 2010.

[24] C. Trattner, I. Hasani-Mavriqi, D. Helic, and H. Leitner. The austrian way of wiki(pedia)!: development of a structured wiki-based encyclopedia within a local austrian context. In *Proceedings of the 6th International Symposium on Wikis and Open Collaboration*, WikiSym '10, pages 1–10, New York, NY, USA, 2010. ACM.

[25] C. Trattner, C. Körner, and D. Helic. Enhancing the navigability of social tagging systems with tag taxonomies. In *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, i-KNOW '11, pages 18:1–18:8, New York, NY, USA, 2011. ACM.

[26] C. Trattner, M. Strohmaier, D. Helic, and K. Andrews. The benefits and limitations of tag clouds as a tool for social navigation. *Technical Report – IICM, Graz University of Technology*, 2011.

9

Improving the Navigability of Tagging Systems with Hierarchically Constructed Resource Lists and Tag Trails

This chapter is based on the contribution *'Improving the Navigability of Tagging Systems with Hierarchically Constructed Resource Lists and Tag Trails"* published in the Journal of Computing and Information Technology in 2011.

This chapter concludes the last part of this dissertation by introducing a generic approach for the construction of resource lists in tagging systems that support efficient navigation of the resources. In detail, the chapter unifies the ideas of the previous chapters to extract hierarchical structures from tagging data automatically and to use these hierarchies for the construction of results lists. Contrary to previous work, the method featured in this chapter is completely generic, i.e. the introduced resource list generation approach could be used to improve the navigability of any tagging system. In a number of experiments based on simulations, we show that the approach generates tag cloud networks which are efficiently navigable.

The original contribution can be found in [22].

9.1 Abstract

Recent research has shown that the navigability of tagging systems leaves much to be desired. In general, it was observed that tagging systems are not navigable if the resource lists of the tagging system are limited to a certain factor k. Hence, in this paper a novel resource list generation approach is introduced that addresses this issue. The proposed approach is based on a hierarchical network model. The paper shows through a number of experiments based on a tagging dataset from a large online encyclopedia system called Austria-Forum, that the new algorithm is able to create tag network structures that are navigable in a efficient manner. Contrary to

previous work, the method featured in this paper is completely generic, i.e. the introduced resource list generation approach could be used to improve the navigability of any tagging system. This work is relevant for researchers interested in navigability of emergent hypertext structures and for engineers seeking to improve the navigability of tagging systems.

9.2 Introduction

With the emergence of modern Web 2.0 hypertext systems such as Flickr, Delicious, CiteULike or LastFM, tagging systems have emerged as an interesting alternative to traditional forms of hypertext navigation and browsing. Tagging systems allow the user to use a free-form vocabulary to annotate resources with the so-called tags [7, 15]. This is done either for semantic reasons (for example, to enrich information items with additional meta data), conversational reasons (for example, for social signaling) [2] or for organizational reasons (for example, to categorize information) [14]. Regardless of why people tag [17, 19, 18], tags are typically visualized as the so-called tag clouds [2]. Basically, a tag cloud is a selection of tags related to a particular resource. Upon clicking on a tag in the tag cloud, a list of resources related to the tag is presented to the user. Thus, in addition to traditional browsing (through a hierarchal taxonomy) and searching (by entering search terms), tags, respectively tag clouds, provide users with a third orthogonal form of navigation within a collection of resources.

In previous work [9, 10, 6], it was observed that the navigability of tagging systems leaves much to be desired. In particular, in [9, 10] we found that the most common resource list generation approach used these days in tagging systems generates network structures which are *per se* unnavigable [9, 10]. The issue is this: Limiting the resource list to a certain factor k, due to interface space restrictions, fragments the bipartite tag network of a tagging system into large isolated network clusters. This renders the network unnavigable from a network-theoretical point of view. However, in [9, 10] we suggested an approach to overcome this issue by applying a simple greedy resource generation strategy. The "trick" is to select, for every click on a particular tag in the tag cloud, the k related resources at random. In common tag cloud algorithms, for every tag click the same result list is generated. Since different resources are selected, this leads to the effect that the tag network becomes connected (even for small values of k) and in theory navigable again. However, as we have shown in [9], this simple strategy does not lead to tag networks which are "good" or even "efficiently" navigable. Therefore, we have investigated in our recent work more sophisticated strategies to generate a k-limited resource list for a particular tag in the tagging system. In [23] we have shown that it is possible, at least in theory, if we apply a hierarchical network model [13] to select the k resources for

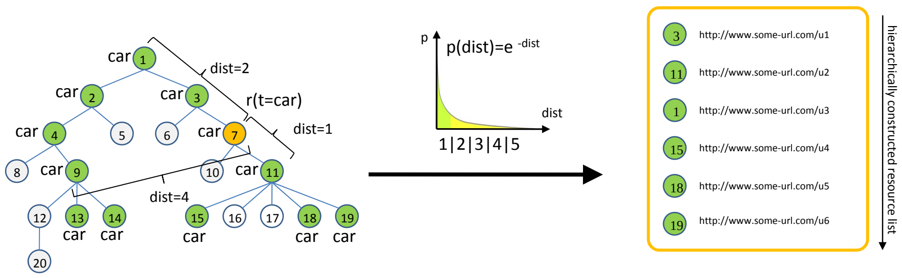


Figure 9.1: Sample resource taxonomy and corresponding hierarchically constructed resource list for tag "car". The green nodes are the resources in the taxonomy that have the tag "car" applied. In the middle of the Figure the resulting probability function is presented and on the right side the generated resource list is shown.

the resource list. The idea is to place the resources in the collection within a hierarchical taxonomy and to use this taxonomy to generate a probability function to select the k resources in the resource list [23]. However as also shown in [23], the approach performs poor, if the resource taxonomy of the system has high branching factors or is poor balanced.

To that end, in this paper we present an enhanced version of the algorithm. Contrary to the approach in [23], the method introduced in this work is able to generate a fixed branched resource taxonomy and corresponding "resource trails" autonomously, i.e. it is independent of any given resource taxonomy. Based on our simulation framework used in [23] we show the high performance of our idea and show that the tag cloud networks generated by this approach, are indeed efficiently navigable.

The paper is structured as follows: In Section 9.3 the hierarchical resource list generation algorithm is presented. In Section 9.4 the dataset used for the experiments is discussed and in Section 9.5 the approach is evaluated. Finally in Section 9.8 the paper is concluded.

9.3 Hierarchical Resource List Construction

The hierarchical resource list generation algorithm is a novel approach for resource list generation in a tagging system [23]. To put it simply, the the approach places the resources into a hierarchical taxonomy and reuses the hierarchy to generate a probability function to select the resources in the tagging system. If the taxonomy provides a constant branching factor b, the emerging tag network is efficiently navigable. The idea for this algorithm was originally derived based upon work by J. Kleinberg [13] who has investigated structural clues of small world networks. Kleinberg showed [13] that if the nodes of a network can be organized into a hierarchy with a constant

Algorithm 8 Hierarchical Resource List Generation Algorithm

INPUT: tag t, resource r, max. resource list size k, resource taxonomy T
OUTPUT: resource list RS
$R(t) \leftarrow$ get all resources $r(t) \backslash r$
$D \leftarrow$ new $HashMap$[new $Array$[]]
for each $r(t)_i \in R(t)$ **do**
    $dist \leftarrow h(r) + h(r(t)_i) - 2h(r, r(t)_i) - 1$
    /* $h(r), h(r(t)_i)$ are the heights of the resource nodes $r, r(t)_i$ in T, $h(r, r(t)_i)$ is the height of the least common ancestor of $r, r(t)_i$ in T */
    $D[dist].add(r(t)_i)$
end for
$j \leftarrow 0$
$RS \leftarrow$ new $Array$[]
while $sizeof(RS) < k$ && $sizeof(RS) < sizeof(D)$ **do**
    $RS[j] \leftarrow D[p_{exp}, p_{uni}]$
    /* p_{exp} is a random number with exponential distribution in the interval $0 \leq x < sizeof(D)$, p_{uni} is a random number with uniform distribution in the interval $0 \leq x < sizeof(D[p_{exp}])$ */
end while
sort RS by $dist$ in descending order
return RS

branching factor b, then such a hierarchy provides a probability distribution for connecting the nodes in the network to generate a network that is then efficiently navigable.

In detail, the algorithm works as follows: For each click on a tag $t(r)$, where r is a resource in the tagging system, the algorithm returns a k-limited resource where the resources $r(t(r))_i$ in the list are selected randomly according to a probability function p that is calculated from a given resource taxonomy T. p is calculated as

$$e^{-dist(r(t(r)), r(t(r))_i)} \tag{9.1}$$

The distance $dist(r(t), r(t)_i)$ is calculated as

$$h(r(t)) + h(r(t)_i) - 2h(r(t), r(t)_i) \tag{9.2}$$

where $h(r(t))$, $h(r(t)_i)$ are the heights of $r(t)$ and $r(t)_i$ in a given resource taxonomy T and where $h(r(t), r(t)_i)$ is the height of the least common ancestor of $r(t)$ and $r(t)_i$ in the resource taxonomy T [23] (see Algorithm 8).

In Algorithm 9.1 an illustrative example of a resource taxonomy and the corresponding hierarchically constructed resource list for the tag "car" is given. Note that the orange node in Algorithm 9.1 represents the resource that is currently viewed by the user. The green nodes are the resources in the taxonomy that have the tag "car" applied. The resulting probability function is presented in the middle of Algorithm 9.1 and the generated resource list is shown on the right side.

9.3.1 Resource Taxonomy Generation Algorithm

To overcome the issue of a given resource taxonomy Algorithm 8 has been extended to a generate a fixed branched resource taxonomy autonomously. In related work, [11] Heymann et al. (see also [4]) describe an algorithm to generate a tag taxonomy from tagging data. The input for the algorithm is the so-called tag similarity graph, i.e. an unweighted graph where each tag is a node in the graph, and two nodes are linked to each other if their similarity is above a predefined similarity threshold. In the simplest case, the threshold is defined by tag overlap, i.e. tags need to share at least one resource to be linked with each other. The second prerequisite for the algorithm is the ranking of nodes in a descending order according to how central the tags are in the tag similarity graph. In particular, this ranking produces a generality order where the most general tags from a dataset are highly ranked. The algorithm starts with the most general tag as the root node of the tree. The algorithm then proceeds by iterating through the generality list. For each tag in the the tree it adds the current processed tag as a child to its most similar tag. [8]

In this work, a similar algorithmic approach is developed. Contrary to the algorithm of Heymann et al. the algorithm is able to generate a fixed branched taxonomy without defining a predefined similarity threshold. In Algorithm 9 the actual algorithm is presented. In words, the algorithm works as follows (see also [25]):

The algorithm takes a tag dataset and the desired taxonomy branching factor as input parameters. Since the algorithm should generate a resource taxonomy with the most general resource of the tagging system as root node and related and less general resources as children, the algorithm calculates in the first step degree centrality for all resource of the supplied tagging dataset and stores the centrality-resource pairs into a map C. Degree centrality was chosen since, on the one hand, it is computed fast, and on the other hand, it was observed in previous research [5] that degree centrality in tagging systems is highly correlated to sophisticated centrality measures such as closeness or betweenness centrality. In the next step, the algorithm sorts the resources in C according to their centrality values in descending order.

Subsequently, the algorithm takes the first element of C (i.e. the most general resource) and sets that resource as the root node of the resource taxonomy. Thereafter, the algorithm starts iterating trough the elements (resources) already present in resource taxonomy. For each resource in the resource taxonomy the algorithm calculates then the most similar resources (see $getMoreLikeThis$). Our algorithm calculates cosine similarity between all co-occurring resources taking also the $tf \cdot idf$ values of the tag concepts into account. Additionally, the function returns only resources that are not already part of the constructed resource taxonomy. The results of this func-

Algorithm 9 Resource Taxonomy Generation Algorithm

```
INPUT: Tag Dataset D, Branching Factor b
OUTPUT: Resource Taxonomy T
C ← new HashMap[]
T ← new Tree[]
for each r_i ∈ F do
    C[r_i] ← calculate degree centrality
end for
sortByValues(C)
/*sort C by values in descending order*/
T[0] ← C[0]
SIM ← new HashMap
for i = 0; i < sizeof(T); i++ do
    /*get all similar resources of T[i] and store the resources as key values and
    the similarity values into SIM*/
    SIM ← getMoreLikeThis(T[i])
    for each r_i ∈ SIM do
        T[r_i] ← T[r_i] · C[r_i]
    end for
    /*sort the resources in SIM by values in descending order*/
    sortByValues(SIM)
    for j = 0; j < sizeof(SIM) and j < b; j++ do
        T[i].append(SIM[i])
    end for
end for
return T
```

tion are stored into a map SIM, with resources as key values and with the provided similarity values as corresponding map values. To account for resource generality we multiply resource similarity values with their corresponding centrality values. The final scores are normalized to fall into the range of [0...1]. After that, the resources in SIM are sorted by the scores in descending order. This procedure ensures that the resources in SIM are not only similar to the currently processed resource but also sorted by their generality values. Thereafter the algorithm appends a maximum of b resources to the currently processed resource. The algorithm stops, if no more similar resources can be found.

Note, due the fixed branching factor b the algorithm does not guarantee that all resources of the tagging dataset are contained in the resulting resource taxonomy. However, as in [23] the probability that one or even more resources are missing is relatively small due to the high number of existing links between the resources of the resource-to-resource network of a given tag dataset. On the other hand, in a tag taxonomy the probability that one concept is missing is significantly higher. The reason for this behavior is the fact that the tag-to-tag network of a tagging system is typically substantially less connected.

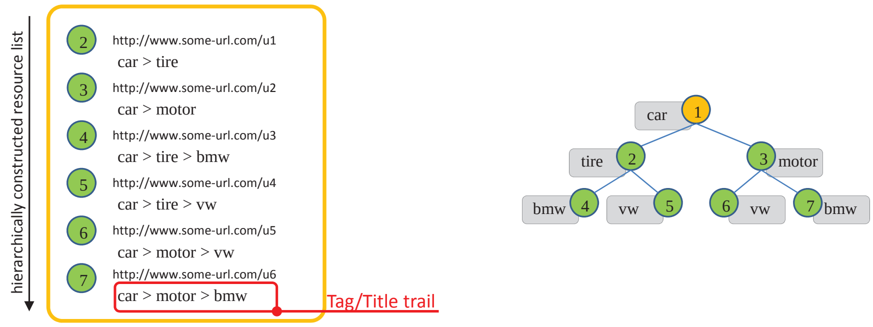
Algorithm 10 Resource Taxonomy Labeling Algorithm

```
INPUT: Resource Taxonomy T, Tag Dataset D
OUTPUT: Tag-resource Taxonomy
COTAGS ← new HashMap[newArray[]]
for i = 0; i < sizeof(T); i++ do
   Ts ← getTags(T[i], D)
   for j = 0; j < sizeof(Ts); j++ do
      cotags ← getCoocTags(Ts[j], D)
      sort(cotags)
      remove all tags from cotags that are not contained in T[i]
      COTAGS[T[i]].add(cotags)
   end for
end for
trails ← new HashSet[]
for each r_i ∈ T do
   /*T is traversed in left-order*/
   pl ← getParentLabels(r_i)
   for each l_j ∈ COTAGS[r_i] do
      if !pl.contains(l_j) then
         if !(trails.contains(pl.toString() + l_j)) then
            T[r_i].applyLabel(pl)
            trails.add(pl.toString() + l_j)
         end if
      end if
      if T[r_i] has no label then
         T[r_i].applyLabel(getTitle(r_i))
      end if
   end for
end for
return T
```

9.3.2 Resource Taxonomy Labeling Algorithm

In order to give the user information about how the resources are structured in the tagging system, tag/title trails are attached as additional information for each resource of the tagging system (see Figure 9.2). In an experiment [21] conducted recently, resource trails were attached to the resources in the result lists of the tagging system. In other words, in [21] we observed that all 24 participants of the experiment were using resource trail information for orientation rather than tag information to navigate the tagging system.

However, since resource trails extracted from a resource taxonomy would be impossible for humans to read, a labeling algorithm is introduced to make the resource taxonomy readable by humans. The basic idea of the algorithm (see Algorithm 8) is it to use tag and title information to label a particular resource in the resource taxonomy and to use the resulting taxonomy to generate tag/title trails which are attached to the resources in the resource list (see Figure 9.2), i.e. we attempt to contextualize the resources.



(a) Sample of a hierarchically constructed resource list with attached tag/title trails

(b) Sample resource taxonomy with attached tag/title labels

Figure 9.2: Sample of a hierarchically constructed resource list with attached tag/title trails (on the left) and corresponding resource taxonomy with applied tag/title labels (on the right). Note, compared to a pure tag taxonomy (see [8] for instance), in a labeled resource taxonomy terms can occur more than once. The orange node in the resource taxonomy (again) denotes the currently viewed resource by the user.

In general it is a labeling algorithm taking a given resource taxonomy and a tagging dataset as input parameters. Tag information is used as label data. The algorithm tries to apply labels to the given resource taxonomy in such a way, that they are uniquely distinguishable and the most descriptive for the given resource. The candidate tags are thereby ranked by the method of tag co-occurrence. However, since it can happen that resources in the resource taxonomy have the same tags in their parent tag trail, due to the lack of available tags in the tagging system, additional meta-data is taken into account. We use title information of the resources as an additional way for differentiation. In words the algorithm works as follows (see also [25]):

In the first step the algorithm calculates, for each resource in the resource taxonomy a list of co-occurring tags of all resource tags and stores this list sorted in descending order into a map. After that, the algorithm traverses the resource taxonomy in left-order. In this loop the actual labeling procedure is performed. In detail, the labeling process looks as follows: For each resource in the resource taxonomy the corresponding co-occurrence vector is consulted and the first label, i.e. the most frequent tag, is tried to be applied to the currently processed resource. If the currently used candidate tag is already part of the tag trail of the currently processed resource (see variable *trails* in Algorithm 9) the next element, i.e. the next frequent tag label is chosen as candidate tag. If no uniquely distinguishable tag trail can be constructed, i.e. the candidate tag label from the co-occurrence vector is already present in the tag trail of the resource additional meta data is considered. We use title information of the currently processed resource for this purpose. Note, since tag and title information can be identical the proposed method is not completely free of collisions. However, to fix this issue one can include additional meta data information or other methods to generate a unique label such as appending an iterative number for each label that occurs more than once. The algorithm stops if all resources of the given resource taxonomy are labeled.

9.4 Dataset

The described experiments in this paper are based on the tag dataset from a system called the Austria-Forum [3, 24]. Basically, the Austria-Forum is a large online encyclopedia similar to Wikipedia providing the user with approximately 180,000 resources related to Austria. In contrast to Wikipedia, Austria-Forum structures articles into a taxonomy and provides an integrated tagging system [24, 20], which allows users to assign tags to resources and to navigate to related resources via tag clouds. As of October 16, 2010 the Austria-Forum tag dataset contains 97,908 tag assignments, 13,314 tags, and 19,430 resources.

9.5 Experiments

In order to evaluate the proposed hierarchical resource list generation approach overall two different experiments were conducted. Since evaluation on the usefulness of the hierarchy creation algorithm was already published in [25], we only present results on the navigability of this approach in this paper.

9.6 Measuring Navigability

In order to evaluate the navigability of the tag networks resulting from the proposed hierarchical resource list generation approach, two different types of tag networks were generated. They all varied in how the the resource lists were calculated. In the following list, we describe the tag networks as they were generated and used for our further experiments:

- **Network $RAND$:** This type of tag network relies on the resource list generation algorithm that returns for a particular tag t a different and randomly sorted k-limited resource list. Contrary to the chronological approach, the resource lists are *not* statically calculated, i.e. for each click on a tag $t(r)$, a different resource list is generated.

- **Network $HIERx$:** This type of tag network relies on the hierarchical resource list generation approach. For our experiments in this paper

Name	k	Nodes	Links	ED	LSCC	NAV
$RAND_{10}$	10	19,430	678,623	4.00	0.99	nav.
$HIER2_{10}$	10	19,430	619,641	4.29	0.99	nav.
$HIER5_{10}$	10	19,430	622,554	3.99	0.99	nav.
$HIER10_{10}$	10	19,430	625,512	4.30	0.99	nav.
$HIER5_{10}$	10	19,430	622,554	3.99	0.99	nav.
$RAND_{50}$	50	19,430	2,191,483	3.87	0.99	nav.
$HIER2_{50}$	50	19,430	2,086,978	4.05	0.99	nav.
$HIER5_{50}$	50	19,430	2,093,926	3.90	0.99	nav.
$HIER10_{50}$	50	19,430	2,097,897	3.86	0.99	nav.

LSCC = Largest Strongly Connected Component, ED = Effective Diameter, NAV = Navigability

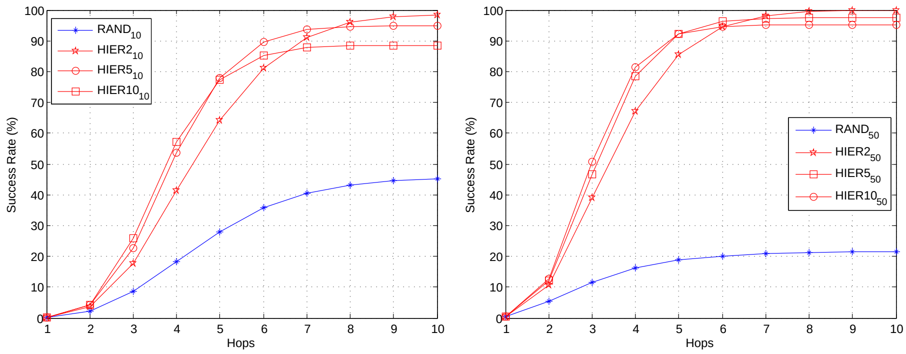
Table 9.1: Tag network statistics. According to Kleinberg [12, 13] networks $RAND$ and $HIERx$ are navigable networks.

three separate tag networks of this type were generated. They all vary in the way in which resource taxonomies were used to generate the resulting tag networks. As input resource taxonomies the three resource taxonomies $Res2$, $Res5$ and $Res10$ were chosen. The resulting networks are called Network $HIER2$, $HIER5$ and $HIER10$.

Chronological, similarity or popularity ranking of the resources was not considered in this evaluation since our previous work in this area [23] showed that they generate network structures which are not navigable.

In order to determine whether the generated tag networks are navigable, network properties such as the size of the largest strongly-connected component ($LSCC$) and the effective diameter (ED) were calculated. From a network-theoretic perspective, Kleinberg [13] showed that a navigable network can be formally defined as a network with a low diameter [16] bounded by $log(n)$, where n is the number of nodes in the network, and an existing giant component, i.e. a strongly connected component containing almost all of the nodes. For that experiment the maximum resource list size k was also varied to $k = 10$ and $k = 50$. This was done to observe whether or not different values of k influence the navigability of the different tag networks. The overall goal of this experiment was to determine whether or not the tagging system relaying on a hierarchical resource list generation algorithm produces tag networks which are more navigable than tag networks generated by a random resource list generation approach.

In Table 9.1, the network statistics of all four tag networks are shown. According to Kleinberg [12, 13] bot networks $RAND$ and $HIER$ are navigable networks.



(a) $k = 10$ **(b)** $k = 50$

Figure 9.3: Success rates of the hierarchical decentralized for tag networks $RAND$ and $HIERx$ and different values of k. As shown, the hierarchically constructed tag networks outperform the random networks most. As also shown, tag network $HIER2$ is most navigable. Regardless of which branching factor, the searcher is able to find nearly 100% of all nodes in this network. According to Kleinberg's definition [12, 13] tag networks $HIER2$ and $HIER5$ are also efficiently navigable network.

9.7 Measuring Efficiency

Since, the previous experiment did not show any difference between the two types of networks an additional simulation was performed. For that purpose we measured the efficiency of the tag network with a hierarchical decentralized searcher [1] as introduced in [8]. As defined by Kleinberg, an efficiently navigable network is a network for which a decentralized searcher exists that is able to navigate to all nodes of the network in $log(n)$ or at least in sub-linear to n time, where n are the number of nodes in the network. In [8] we have introduced a searcher that is able to search a tagging system in $log(n)$. However, contrary to the searcher in [8], the searcher in this work uses as background knowledge the resource taxonomy which was utilized to generate the tag network (see Algorithm 11). Additionally, the searcher in this work is able to walk along a directed tag network. In [8], it was limited to a bipartite tag network. In Algorithm 11, the pseudo code of the implemented searcher is presented. Note that the searcher is using as input parameters a directed resource-resource tag network, a resource taxonomy, a start and target node and a maximum number of hops parameter that defines how many resources the searcher should at maximum visit before giving up. For an input taxonomy the searcher is taking the corresponding resource taxonomy, i.e. for Network $HIER2$ resource taxonomy $RES2$ is taken, for Network $HIER5$ resource taxonomy $RES5$ is taken, for Network $HIER10$ resource taxonomy $RES10$ is taken and for network $RAND$ a

Algorithm 11 Hierarchical Decentralized Searcher [8]

INPUT: resource resource graph R, resource taxonomy T, start node v, target node w, max hops $hops_{max}$
$hops \leftarrow 0$
while $v! = w$ **do**
  if $++hops >= hops_{max}$ **then**
    break
  end if
  $R(v) \leftarrow$ get all resources from $v \in R$
  $dist_{min} \leftarrow \infty$
  for each $r_i \in R(v)$ **do**
    $dist \leftarrow h(r_i, T) + h(v, T) - 2h(r_i, v, T) - 1$
    if $dist < dist_{min}$ **then**
      $dist_{min} \leftarrow dist$
      $v \leftarrow r_i$
    end if
  end for
end while

random resource taxonomy was generated.

In order to acquire statistically significant results, 100,000 random searches (with a maximum of 10 hops) for each of the networks were performed. The start and target nodes were selected uniform at random. For the experiment only resource pairs were considered for which a path was present in the network. If the target node could not be found in at least 10 hops or the searcher was caught in a cycle (we did not recover the searcher in that case) this was counted as an error. It is important to note that both searcher were given the exact same start and target nodes for all four networks.

In Figure 9.3 we present the success rate plots of the hierarchical decentralized searcher for tag networks $RAND$ and $HIERx$ and different values of k. As shown, the hierarchically constructed tag networks outperform the random networks significantly. As also shown in Figure 9.3, tag network $HIER2$ is most navigable. Regardless of which branching factor, the searcher is able to find nearly 100% of all nodes in this network. According to Kleinberg's definition [12, 13] tag networks $HIER2$ and $HIER5$ are also efficiently navigable network.

9.8 Conclusions

In this paper, a novel approach for resource list generation for tagging systems was presented. It continues our work on the navigability of social tagging systems and presented a resource list generation approach that is based on a hierarchical network model. A number of experiments showed that the approach is able to generate tag network structures which are efficiently navigable. Contrary to previous work, the proposed approach is completely generic, i.e. the introduced hierarchical resource list generation approach could be used to improve the navigability of any tagging system.

References

[1] L. Adamic and E. Adar. How to search a social network. *Social Networks*, 27(3):187 – 203, 2005.

[2] M. Ames and M. Naaman. Why we tag: Motivations for annotation in mobile and online media. In *Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI 2007)*, CHI'07, pages 971–980, USA, NY, 2007. ACM.

[3] Austria-Forum. Das Oesterreichische Wissensnetz. http://wwww.austria-lexikon.at, 2011. [Online; accessed 2011-03-01].

[4] D. Benz, A. Hotho, and G. Stumme. Semantics made by you and me: Self-emerging ontologies can capture the diversity of shared knowledge. In *Proc. of the 2nd Web Science Conference (WebSci10)*, Raleigh, NC, USA, 2010.

[5] D. Benz, C. Körner, A. Hotho, G. Stumme, and M. Strohmaier. One tag to bind them all : Measuring term abstractness in social metadata. In *Proceedings of the 8th Extended Semantic Web Conference (ESWC 2011)*, Heraklion, Crete, May 2011.

[6] E. H. Chi and T. Mytkowicz. Understanding the efficiency of social tagging systems using information theory. In *HT '08: Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*, pages 81–88, New York, NY, USA, 2008. ACM.

[7] T. Hammond, T. Hannay, B. Lund, and J. Scott. Social bookmarking tools (i): A general review. *D-Lib Magazine*, 11(4), 2005.

[8] D. Helic, M. Strohmaier, C. Trattner, M. Muhr, and K. Lermann. Pragmatic evaluation of folksonomies. In *Proc. of the 21st International World Wide Web conference, WWW '11*, New York, NY, USA, 2011. ACM.

[9] D. Helic, C. Trattner, M. Strohmaier, and K. Andrews. On the navigability of social tagging systems. In *Proc. of 2010 IEEE International Conference on Social Computing*, pages 161–168, Los Alamitos, CA, USA, 2010. IEEE Computer Society.

[10] D. Helic, C. Trattner, M. Strohmaier, and K. Andrews. Are tag clouds useful for navigation? a network-theoretic analysis. *International Journal of Social Computing and Cyber-Physical Systems*, 1(1):33–55", 2011.

[11] P. Heymann and H. Garcia-Molina. Collaborative creation of communal hierarchical taxonomies in social tagging systems. Technical Report 2006-10, Stanford InfoLab, April 2006.

[12] J. Kleinberg. The small-world phenomenon: An algorithmic perspective. In *Proceedings of the 32nd ACM Symposium on Theory of Computing*, 2000.

[13] J. M. Kleinberg. Small-world phenomena and the dynamics of information. In *Advances in Neural Information Processing Systems (NIPS) 14*, page 2001. MIT Press, 2001.

[14] C. Körner, R. Kern, H. P. Grahsl, and M. Strohmaier. Of categorizers and describers: An evaluation of quantitative measures for tagging motivation. In *21st ACM SIGWEB Conference on Hypertext and Hypermedia (HT 2010), Toronto, Canada, ACM*, Toronto, Canada, June 2010.

[15] C. Marlow, M. Naaman, D. Boyd, and M. Davis. Ht06, tagging paper, taxonomy, flickr, academic article, to read. In *Proc. Seventeenth Conference on Hypertext and Hypermedia (Hypertext 2006)*, HT'06, pages 31–40, USA, NY, 2006. ACM.

[16] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[17] O. Nov and C. Ye. Why do people tag?: motivations for photo tagging. *Commun. ACM*, 53:128–131, July 2010.

[18] M. Strohmaier. Purpose tagging - capturing user intent to assist goal-oriented social search. In *SSM'08 Workshop on Search in Social Media, in conjunction with CIKM'08*, USA, Napa Valley, 2008.

[19] M. Strohmaier, C. Koerner, and R. Kern. Why do users tag? Detecting users' motivation for tagging in social tagging systems. In *International AAAI Conference on Weblogs and Social Media (ICWSM2010), Washington, DC, USA, May 23-26*, 2010.

[20] C. Trattner. Querycloud: Automatically linking related documents via search query (tag) clouds. In *Proceedings of IADIS International Conference WWW/Internet 2010*, ICWI'10. IADIS, 2010.

[21] C. Trattner. Improving tag cloud navigability: A comparative study. In *Proc. of the 33th International Conference on Information Technology Interfaces*, Cavtat, Dubrovnik, Croatia, 2011. IEEE.

[22] C. Trattner. Improving the navigability of tagging systems with hierarchically constructed resource lists and tag trails. *CIT*, 19(3):155–167, 2011.

[23] C. Trattner, H. Denis, and M. Strohmaier. On the construction of efficiently navigable tag clouds using knowledge from structured web content. *Journal of Universal Computer Science*, 2011.

[24] C. Trattner, I. Hasani-Mavriqi, D. Helic, and H. Leitner. The austrian way of wiki(pedia)!: development of a structured wiki-based encyclopedia within a local austrian context. In *Proceedings of the 6th International Symposium on Wikis and Open Collaboration*, WikiSym '10, pages 1–10, New York, NY, USA, 2010. ACM.

[25] C. Trattner, C. Körner, and D. Helic. Enhancing the navigability of social tagging systems with tag taxonomies. In *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, i-KNOW '11, pages 18:1–18:8, New York, NY, USA, 2011. ACM.

Part V

Summary and Outlook

10

Conclusions and Future Directions

This chapter summarizes and concludes the dissertation. We first list the main contributions of this thesis and provide then the answers to the research questions as proposed in the introductory part of this thesis. Last but not least, we discuss the limitations of our work and propose future directions of research in the field of tag-based search and navigation in tagging systems.

10.1 Summary of Contributions

The novelty of this work lies in the idea of reviewing social tags and corresponding browsing constructs from a navigational perspective. While related work has studied the utility of tags mostly from the visual point of view or a information-theoretic perspective, to the best of our knowledge this is the first work that extensively studied the extent to which tags and corresponding constructs are useful for *efficiently* searching and *navigating* to the resources of a tagging system.

The following list summarizes the contributions of this dissertation:

1. The review of the utility of tags for the task of search and efficient navigation in tagging systems (see Part II).
2. The navigational review of tags compared to other tag-alike meta-data structures such as keywords and search query terms (see Part III).
3. The introduction of a number of new approaches that support efficient tag-based navigation in tagging systems (see Part IV).

10.2 Answers to Research Questions

At the beginning of this dissertation we set out a number of research questions that we were interested in answering. The following section summarizes the answers to these questions.

Research Question 1

To what extent are tags/tag clouds useful for navigation?

The first research question we focused on in this dissertation was the issue to what extent tags are useful for navigation. Since related work has only partly answered this question from an information retrieval perspective and one single tag dataset, we examined this question from a network-theoretic perspective and by using a number of tag datasets. In general, we showed that tags are theoretically useful for navigation. By modeling tagging systems as tag-resource networks, we showed that tagging systems have power-law qualities, a giant component that includes over 99% of the resources and an effective diameter that is bound by $log(n)$, where n are the number of resources in a tagging system. By utilizing the network-theoretic notations introduced by J. Kleinberg, we could show that tags are an efficient source for navigating the resources of an information system.

While studying the utility of tags from a navigational perspective, we also rose the question as to what extent tag clouds are useful for navigation. Since tags are typically displayed as tag clouds and limited in their size due to interface limitations, we analyzed the utility of tag clouds for the task of navigation in tagging systems. In that context we limited the number of tags displayed in the tag cloud to a factor N as well as the number of resources presented in the result list to a factor k. With navigability measures from the domain of network-theory we could show that for $N > 20$ the efficiency of tag clouds for navigation is not impaired. Contrary to this, if we limited the pagination factor k to sizes of $5, 10, 20$ or 30 we could show that the resulting tag-resource network was not navigable due to the fact of a missing giant component that connected almost all resources in the tagging system with each other.

Research Question 2

To what extent are tags/tag clouds useful for search?

After studying the utility of tags for navigation, we were interested on the usefulness of tags for task of search in tag-based information systems. While related work showed that tags can enhance the information retrieval properties of an tag-based information systems, or has shown that tag-clouds might be useful for summarizing search results, we were interested studying the extent to which tags/tag clouds are useful for search to enhance the performance and users' satisfaction. To that end, we explored the differences between three search interfaces under a controlled user study. The interfaces explored in the study included a search-only interface that plays the role of a baseline and two other search interfaces: a tag cloud based search interface and a faceted tag cloud interface. In order to evaluate the interfaces we utilized two well-known information seeking strategies typically performed by users to search in an information system: look-up search and exploratory search. In a number of experiments with 24 users and a within-subject design, we could show that the two tag-based search interfaces performed better than the baseline in both user satisfaction and performance. However, as our results also demonstrated, the differences between the two tag-cloud based search interfaces are not so clear. While users performed in the tag cloud search interface best and also preferred it most over all other interfaces, they would recommend the faceted tag cloud interface.

Research Question 3

> *To what extent are tags/tag clouds more useful/efficient for search/ navigation than other tag-alike meta-data such as keywords or search query-terms?*

Another question which we were interested in was the question to what extent tags are more useful for navigation than tag-alike meta-data such as keywords or query-tags. Since tags are very related to the notation of keywords and since related research has shown that tags are in their structure comparable to so-called query tags harvested from search query logs, we were interested to study the navigational differences and similarities of tags compared to these tag-alike meta-data structures. To that end we conducted two studies comparing tags from a navigational perspective with query terms and tags with keywords.

The first study introduced QueryCloud, a tool that harvests query terms in an online encyclopedia system called Austria-Forum to generate query tag clouds for the purpose to link related content within the system automatically. On a theoretical and empirical level we could show that QueryCloud out-performs the integrated tagging system of the Austria-Forum by generating tag clouds that are more efficiently navigable than the tag clouds which are based on tags generated by real-users of the system. Additionally to this, we could show in a user study that query tags generated by QueryCloud are almost to the same degree relevant for the user of a given Web page as tags generated by real users.

In the second study, we explored the navigational differences between broad (tag-based) and narrow (keyword-based) folksonomies in social hypertext systems. We studied both kinds of folksonomies on a dataset provided by Mendeley - a collaborative platform where users can annotate and organize scientific articles with tags and keywords. In a variety of experiments based on information- and network-theory we could show that broad folksonomies (=tag-based folksonomies) are more efficient for navigation than narrow folksonomies (=keyword-based folksonomies).

Research Question 4

To what extent can we build better tag-based browsing constructs that support better search/navigation in tagging systems?

Since our research on tag-based browsing showed that tag clouds are limited in their functionality to support efficient navigation of the resources of a tagging system mainly because of the so-called pagination effect, we were interested in developing better tag-based browsing constructs that support more efficient navigation in tagging systems than currently available approaches. To that end, we presented three studies that introduced overall three novel approaches for the construction of tag clouds and tag hierarchies that support better navigation in tagging systems.

The first study introduced the notation of hierarchically constructed tag clouds and resource lists. In this study we showed on a network-theoretic and empirical level that we can construct tag clouds that are more navigable than current available approaches by utilizing a given hierarchical resource structure of the system.

In the second study we introduced the approach of the so-called tag-resource taxonomies which on the one hand address the problem to create fixed branched and semantically sound resources taxonomies automatically from tagging data and which on the other hand performs better for navigating the resources of a tagging system than regular tag taxonomies. This was shown again on a theoretic and empirical level.

The third and last study combined the two previous approaches and introduced the idea of hierarchically constructed resource lists with tag trails. In a number of experiments based on simulations we could show that the approach generates network structures which are indeed efficiently navigable.

10.3 Limitations and Future Directions

Although this dissertation presents a large number of results on the usefulness of tags or tag clouds for the task search and navigation in tagging systems and includes approaches such as hierarchically constructed tag clouds or tag-resource hierarchies that enhance the navigability of tagging systems, the work in this area is still at an early stage. In this section we aim to focus on the limitations of our research as well as on future directions of our work.

10.3.1 Limitations

Despite the high number of experiments performed in this dissertation to shed light onto the usefulness of tags for the task of search and navigation in tag-based information systems we have to acknowledge some limitations of our work. The following two sections discuss two possible points of concern in our dissertation.

To what extent are hierarchically created tag clouds more useful than other tag cloud construction approaches?

One of the potential criticism regarding the approaches introduced in Part IV could be that we only focused on comparing our method with the most popular method for constructing tag clouds in tagging systems – formally introduced as the TopN tag cloud construction algorithm with reverse chronologically sorted result lists. As presented in Chapter 2 and in the related work sections in Part IV, there are a large volume of tag cloud construction algorithms that display the tags in the tag cloud in a clustering or other similarity based manner. Since most of these algorithms are based on popularity or generality measures, we believe that the effect on the navigability of tagging systems is the same as with the popular TopN tag cloud calculation algorithm. However, to justify this assumption, further research in this area is needed.

To what extent is it justified to model tag-based navigation with hierarchical decentralized search?

Another critical point of our dissertation might be that we utilized for our performance analysis a tag-based navigational model which is grounded on Adamic's hierarchical decentralized search approach and did not perform large-scale user studies to confirm our theoretical assumptions. Even if we presented a user study in Chapter 7 confirming our theoretical findings from the same chapter, we recently conducted a large-scale study comparing real user click data with our hierarchical decentralized search model. The paper discussing the results of this study is included in Appendix A and reflects the differences and similarities of hierarchical decentralizes search and human navigation in information networks. As shown, the paper reviews the navigational behavior of users on Wikipedia. Nevertheless, and even if we do not explicitly discussing navigational behavior of users in tagging systems, we believe that the paper presents well the extent to which it is justified to utilize a hierarchical decentralized search routine to simulate human navigational behavior in information networks.

10.3.2 Future Work

However, although this dissertation at hand has presented a large volume of studies in the field of tag-based search and navigation, research in this area is still inconclusive. In the following section a list of potential research questions is presented we would like to study in one of our future works:

- To what extent are tag hierarchies more useful for navigation than tag clouds?
- Are clustered tag cloud approaches more useful for navigation than traditional tag clouds?
- To what extent are tags more efficient for navigation than other kinds of meta-data such as for instance named-entities?
- To what extent are tags more useful in search interfaces than query suggestion mechanisms?
- To what extent is the navigability of tagging systems influenced by the tagging motivations of their users?
- To what extent can we model tag-based navigation with probabilistic models such as Markov-Chains? Are these models a better proxy for simulating tag-based navigation in tagging systems than using hierarchical decentralized search?

Bibliography

[1] P. , G. Koutrika, and H. Garcia-Molina. Can social bookmarking improve web search? *Proceedings of the international conference on Web search and web data mining*, pages 195–206, 2008.

[2] L. Adamic and E. Adar. How to search a social network. *Social Networks*, 27(3):187 – 203, 2005.

[3] L. A. Adamic, R. M. Lukose, A. R. Puniyani, and B. A. Huberman. Search in power-law networks. *Physical Review E*, 64(4):046135 1–8, Sep 2001.

[4] E. Adar, J. Teevan, and S. T. Dumais. Large scale analysis of web revisitation patterns. In *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, CHI '08, pages 1197–1206, New York, NY, USA, 2008. ACM.

[5] S. Alexaki, V. Christophides, G. Karvounarakis, D. Plexousakis, and K. Tolle. The ics-forth rdfsuite: Managing voluminous rdf description bases. In *SemWeb*, 2001.

[6] M. Ames and M. Naaman. Why we tag: Motivations for annotation in mobile and online media. In *Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI 2007)*, CHI'07, pages 971–980, USA, NY, 2007. ACM.

[7] A. Ammari and V. Zharkova. Combining tag cloud learning with svm classification to achieve intelligent search for relevant blog articles. In *In Proceedings of the 1st International Workshop on Mining Social Media*, 2009.

[8] I. Antonellis, H. Garcia-Molina, and J. Karim. Tagging with queries: How and why? In *WSDM (Late Breaking-Results)*, 2009.

[9] I. Antonellis, J. Karim, and H. Garcia-Molina. Navigating the web with query tags. Technical report, Stanford University, 2010.

[10] K. Aouiche, D. Lemire, and R. Godin. Web 2.0 OLAP: From Data Cubes to Tag Clouds. *4th International Conference, WEBIST 2008*, 18, 2008.

[11] Austria-Forum. Das Oesterreichische Wissensnetz. http://wwww.austria-lexikon.at, 2011. [Online; accessed 2011-03-01].

[12] L. B. Baltussen. *Barbarians versus gatekeepers ? Tagging as a way of defining the ememrgent living archive paradigm.* PhD thesis, 2010.

[13] S. Bao, G. Xue, X. Wu, Y. Yu, B. Fei, and Z. Su. Optimizing web search using social annotations. In *Proceedings of the 16th international conference on World Wide Web - WWW '07*, page 501, New York, New York, USA, May 2007. ACM Press.

[14] S. Bateman, C. Gutwin, and M. Nacenta. Seeing things in the clouds: the effect of visual features on tag cloud selections. *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*, pages 193–202, 2008.

[15] D. Benz, A. Hotho, R. Jäschke, B. Krause, and G. Stumme. Query logs as folksonomies. *Datenbank-Spektrum*, 10(1):15–24, 2010.

[16] D. Benz, A. Hotho, and G. Stumme. Semantics made by you and me: Self-emerging ontologies can capture the diversity of shared knowledge. In *Proc. of the 2nd Web Science Conference (WebSci10)*, Raleigh, NC, USA, 2010.

[17] D. Benz, C. Körner, A. Hotho, G. Stumme, and M. Strohmaier. One tag to bind them all : Measuring term abstractness in social metadata. In *Proceedings of the 8th Extended Semantic Web Conference (ESWC 2011)*, Heraklion, Crete, May 2011.

[18] M. Bilenko and R. W. White. Mining the search trails of surfing crowds: identifying relevant websites from user activity. In *Proceedings of the 17th international conference on World Wide Web*, WWW '08, pages 51–60, New York, NY, USA, 2008. ACM.

[19] K. Bischoff, C. S. Firan, C. Kadar, W. Nejdl, and R. Paiu. Automatically identifying tag types. In *Proceedings of the 5th International Conference on Advanced Data Mining and Applications*, ADMA '09, pages 31–42, Berlin, Heidelberg, 2009. Springer-Verlag.

[20] K. Bischoff, C. S. Firan, W. Nejdl, and R. Paiu. Can all tags be used for search? In *Proceedings of the 17th ACM conference on Information and knowledge management*, CIKM '08, pages 193–202, New York, NY, USA, 2008. ACM.

[21] M. Boguñá, D. Krioukov, and K. C. Claffy. Navigability of complex networks. *Nature Physics*, 5:74–80, Jan. 2009.

[22] M. Boguñá, F. Papadopoulos, and D. Krioukov. Sustaining the Internet with hyperbolic mapping. *Nature Communications*, 1:62, Sept. 2010.

[23] D. Böhnstedt, L. Lehmann, C. Rensing, and R. Steinmetz. Automatic identification of tag types in a resource-based learning scenario. In *Proceedings of the 6th European conference on Technology enhanced learning: towards ubiquitous learning*, EC-TEL'11, pages 57–70, Berlin, Heidelberg, 2011. Springer-Verlag.

[24] B. Bollobás and F. R. K. Chung. The diameter of a cycle plus a random matching. *SIAM J. Discret. Math.*, 1(3):328–333, 1988.

[25] B. Bollobás and W. F. de la Vega. The diameter of random regular graphs. *Combinatorica*, 2(2):125–134, 1982.

[26] K. Byström and K. Järvelin. Task complexity affects information seeking and use. *Inf. Process. Manage.*, 31:191–213, March 1995.

[27] K. Candan, L. Di Caro, and M. Sapino. Creating tag hierarchies for effective navigation in social media. *SSM '08: Proceeding of the 2008 ACM workshop on Search in social media*, pages 75–82, 2008.

[28] I. Cantador, I. Konstas, and J. M. Jose. Categorising social tags to improve folksonomy-based recommendations. *Web Semantics: Science, Services and Agents on the World Wide Web*, 9(1), 2011.

[29] M. J. Carman, M. Baillie, R. Gwadera, and F. Crestani. A statistical comparison of tag and query logs. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '09, pages 123–130, New York, NY, USA, 2009. ACM.

[30] C. Cattuto, D. Benz, A. Hotho, and G. Stumme. Semantic grounding of tag relatedness in social bookmarking systems. In *Proc. of International Semantic Web Conference 2008*, volume 5318 of *LNAI*, pages 615–631, Heidelberg, 2008. Springer.

[31] C. Cattuto, C. Schmitz, A. Baldassarri, V. D. P. Servedio, V. Loreto, A. Hotho, M. Grahl, and G. Stumme. Network properties of folksonomies. *AI Commun.*, 20(4):245–262, 2007.

[32] E. H. Chi and T. Mytkowicz. Understanding the efficiency of social tagging systems using information theory. In *HT '08: Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*, pages 81–88, New York, NY, USA, 2008. ACM.

[33] A. Clauset, C. Moore, and M. E. J. Newman. Hierarchical structure and the prediction of missing links in networks. *Nature*, 453(7191):98–101, 2008.

[34] A. Clauset, C. R. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM Rev.*, 51:661–703, November 2009.

[35] Delicious. Delicious – About. http://delicious.com/about, 2012. [Online; accessed 2012-05-21].

[36] K. Dellschaft and S. Staab. On how to perform a gold standard based evaluation of ontology learning. In *Proceedings of ISWC-2006 International Semantic Web Conference*, Athens, GA, USA, November 2006. Springer, LNCS.

[37] I. Dhillon, J. Fan, and Y. Guan. Efficient clustering of very large document collections. In R. Grossman, C. Kamath, and R. Naburu, editors, *Data Mining for Scientific and Engineering Applications*. Kluwer Academic Publishers, 2001.

[38] A. Dieberger. Supporting social navigation on the world wide web. *Int. J. Hum.-Comput. Stud.*, 46(6):805–825, 1997.

[39] A. Diriye, A. Blandford, and A. Tombros. When is system support effective? In *Proceedings of the third symposium on Information interaction in context*, IIiX '10, pages 55–64, New York, NY, USA, 2010. ACM.

[40] P. A. Dmitriev, N. Eiron, M. Fontoura, and E. Shekita. Using annotations in enterprise search. In *Proceedings of the 15th international conference on World Wide Web*, WWW '06, pages 811–817, New York, NY, USA, 2006. ACM.

[41] T. Eda, T. Uchiyama, T. Uchiyama, and M. Yoshikawa. Signaling emotion in tagclouds. In *WWW '09: Proceedings of the 18th international conference on World wide web*, pages 1199–1200, New York, NY, USA, 2009. ACM.

[42] P. Erdos and A. Renyi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5:17–61, 1960.

[43] B. J. J. Frey and D. Dueck. Clustering by passing messages between data points. *Science*, 315(5814):972–976, January 2007.

[44] W.-T. Fu and P. Pirolli. Snif-act: a cognitive model of user navigation on the world wide web. *Hum.-Comput. Interact.*, 22(4):355–412, 2007.

[45] G. W. Furnas, C. Fake, L. von Ahn, J. Schachter, S. Golder, K. Fox, M. Davis, C. Marlow, and M. Naaman. *Why do tagging systems work?* ACM Press, New York, New York, USA, Apr. 2006.

[46] D. F. Gleich, P. G. Constantine, A. D. Flaxman, and A. Gunawardana. Tracking the random surfer: empirically measured teleportation parameters in pagerank. In *Proceedings of the 19th international conference on World wide web*, WWW '10, pages 381–390, New York, NY, USA, 2010. ACM.

[47] S. A. Golder and B. A. Huberman. The structure of collaborative tagging systems. *Journal of Information Science*, 32(2):198–208, 2006.

[48] H. Halpin, V. Robu, and H. Shepherd. The complex dynamics of collaborative tagging. In *Proceedings of the 16th international conference on World Wide Web*, WWW '07, pages 211–220, New York, NY, USA, 2007. ACM.

[49] M. J. Halvey and M. T. Keane. An assessment of tag presentation techniques. In *Proceedings of the 16th international conference on World Wide Web*, WWW '07, pages 1313–1314, New York, NY, USA, 2007. ACM.

[50] T. Hammond, T. Hannay, B. Lund, and J. Scott. Social bookmarking tools (i): A general review. *D-Lib Magazine*, 11(4), 2005.

[51] Y. Hassan-Montero and V. Herrero-Solana. Improving tag-clouds as visual information retrieval interfaces. In *InScit2006: International Conference on Multidisciplinary Information Sciences and Technologies*, 2006.

[52] Y. Hassan-Montero and V. Herrero-Solana. Improving tag-clouds as visual information retrieval interfaces. In *InScit2006: International Conference on Multidisciplinary Information Sciences and Technologies*, 2006.

[53] M. A. Hearst. *Search User Interfaces*. Cambridge University Press, 1 edition, 2009.

[54] M. A. Hearst and D. Rosner. Tag clouds: Data analysis tool or social signaller? In *HICSS '08: Proceedings of the Proceedings of the 41st Annual Hawaii International Conference on System Sciences*, Washington, DC, USA, 2008. IEEE Computer Society.

[55] M. Heckner, M. Heilemann, and C. Wolff. Personal information management vs. resource sharing: Towards a model of information behaviour in social tagging systems. In *Int'l AAAI Conference on Weblogs and Social Media (ICWSM)*, San Jose, CA, USA, May 2009.

[56] D. Helic, , C. Körner, M. Granitzer, M. Strohmaier, and C. Trattner. Navigational efficiency of broad vs. narrow folksonomies. In *Proceedings of the 23rd ACM Conference on Hypertext and Social Media*, HT '12, New York, NY, USA, 2012. ACM.

[57] D. Helic and M. Strohmaier. Building directories for social tagging systems. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, CIKM '11, pages 525–534, New York, NY, USA, 2011. ACM.

[58] D. Helic, M. Strohmaier, C. Trattner, M. Muhr, and K. Lermann. Pragmatic evaluation of folksonomies. In *Proc. of the 21st International World Wide Web conference, WWW '11*, New York, NY, USA, 2011. ACM.

[59] D. Helic, C. Trattner, M. Strohmaier, and K. Andrews. On the navigability of social tagging systems. In *Proc. of 2010 IEEE International Conference on Social Computing*, pages 161–168, Los Alamitos, CA, USA, 2010. IEEE Computer Society.

[60] D. Helic, C. Trattner, M. Strohmaier, and K. Andrews. Are tag clouds useful for navigation? a network-theoretic analysis. *International Journal of Social Computing and Cyber-Physical Systems*, 1(1):33–55", 2011.

[61] D. Helic, C. Trattner, M. Strohmaier, and K. Andrews. Are tag clouds useful for navigation? a network-theoretic analysis. *International Journal of Social Computing and Cyber-Physical Systems*, 2011.

[62] P. Heymann and H. Garcia-Molina. Collaborative creation of communal hierarchical taxonomies in social tagging systems. Technical Report 2006-10, Stanford InfoLab, April 2006.

[63] P. Heymann, A. Paepcke, and H. Garcia-Molina. Tagging human knowledge. In *Proceedings of the third ACM international conference on Web search and data mining*, WSDM '10, pages 51–60, New York, NY, USA, 2010. ACM.

[64] P. Heymann, A. Paepcke, and H. Garcia-Molina. Tagging human knowledge. In *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, pages 51–61, New York, NY, USA, 2010. ACM.

[65] A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. Bibsonomy: A social bookmark and publication sharing system. In *Proceedings of the Conceptual Structures Tool Interoperability Workshop at the 14th International Conference on Conceptual Structures*, pages 87–102, 2006.

[66] A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. Folkrank: A ranking algorithm for folksonomies. In *Proc. FGIR 2006*, 2006.

[67] Y. Kammerer, R. Nairn, P. Pirolli, and E. H. Chi. Signpost from the masses: learning effects in an exploratory social tag search browser. In *Proceedings of the 27th international conference on Human factors in computing systems*, CHI '09, pages 625–634, New York, NY, USA, 2009. ACM.

[68] O. Kaser and D. Lemire. Tag-Cloud Drawing: Algorithms for Cloud Visualization. *Proceedings of Tagging and Metadata for Social Information Organization (WWW 2007)*, 2007.

[69] D. Kelly. Methods for evaluating interactive information retrieval systems with users. *Found. Trends Inf. Retr.*, 3:1–224, January 2009.

[70] C.-C. Kiu and E. Tsui. Taxofolk: A hybrid taxonomy-folksonomy structure for knowledge classification and navigation. *Expert Systems with Applications*, 38(5):6049 – 6058, 2011.

[71] J. Kleinberg. The small-world phenomenon: An algorithmic perspective. In *Proceedings of the 32nd ACM Symposium on Theory of Computing*, 2000.

[72] J. Kleinberg. Complex networks and decentralized search algorithms. In *International Congress of Mathematicians (ICM)*, pages 1019–1044, Zürich, Switzerland, 2006. European Mathematical Society Publishing House.

[73] J. M. Kleinberg. Navigation in a small world. *Nature*, 406(6798), August 2000.

[74] J. M. Kleinberg. Small-world phenomena and the dynamics of information. In *Advances in Neural Information Processing Systems (NIPS) 14*, page 2001. MIT Press, 2001.

[75] J. Kleinfeld. Could it be a big world after all? The six degrees of separation myth. *Society*, April 2002.

[76] K. Knautz, S. Soubusta, and W. G. Stock. Tag clusters as information retrieval interfaces. In *Proceedings of the 2010 43rd Hawaii International Conference on System Sciences*, HICSS '10, pages 1–10, Washington, DC, USA, 2010. IEEE Computer Society.

[77] K. Knautz, S. Soubusta, and W. G. Stock. Tag clusters as information retrieval interfaces. In *HICSS*, pages 1–10, 2010.

[78] M. Kochen, editor. *The Small World.* Ablex, Norwood, NJ, 1989.

[79] C. Körner. Aspects of user motivation in social tagging systems. *Dissertation, Graz University of Technology*, 2012.

[80] C. Körner, D. Benz, A. Hotho, M. Strohmaier, and G. Stumme. Stop thinking, start tagging: Tag semantics arise from collaborative verbosity. In *19th International World Wide Web Conference (WWW2010), Raleigh, NC, USA, April 26-30, ACM*, 2010.

[81] C. Körner, R. Kern, H. P. Grahsl, and M. Strohmaier. Of categorizers and describers: An evaluation of quantitative measures for tagging motivation. In *21st ACM SIGWEB Conference on Hypertext and Hypermedia (HT 2010), Toronto, Canada, ACM*, Toronto, Canada, June 2010.

[82] G. Koutrika, Z. M. Zadeh, and H. Garcia-Molina. Data clouds: summarizing keyword search results over structured data. In *EDBT*, pages 391–402, 2009.

[83] B. Krause, R. Jäschke, A. Hotho, and G. Stumme. Logsonomy - social information retrieval with logdata. In *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*, HT '08, pages 157–166, New York, NY, USA, 2008. ACM.

[84] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá. Hyperbolic geometry of complex networks. *Phys. Rev. E*, 82(3):036106, Sep 2010.

[85] B. Y.-L. Kuo, T. Hentrich, B. M. . Good, and M. D. Wilkinson. Tag clouds for summarizing web search results. In *Proceedings of the 16th international conference on World Wide Web*, WWW '07, pages 1203–1204, New York, NY, USA, 2007. ACM.

[86] J. R. Landis and G. G. Koch. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):pp. 159–174, 1977.

[87] S. Lattanzi, A. Panconesi, and D. Sivakumar. Milgram-routing in social networks. In *Proceedings of the 20th international conference on World wide web*, WWW '11, pages 725–734, New York, NY, USA, 2011. ACM.

[88] J. Leskovec, J. Kleinberg, and C. Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. In *KDD*

'05: Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining, pages 177–187, New York, NY, USA, 2005. ACM.

[89] N. Levine. Strongly connected sets in topology. *The American Mathematical Monthly*, 72(10):pp. 1098–1101, 1965.

[90] R. Li, S. Bao, Y. Yu, B. Fei, and Z. Su. Towards effective browsing of large scale social annotations. *Proceedings of the 16th international conference on World Wide Web*, page 952, 2007.

[91] D. Liben-Nowell and J. Kleinberg. The link prediction problem for social networks. In *Proceedings of the twelfth international conference on Information and knowledge management*, CIKM '03, pages 556–559, New York, NY, USA, 2003. ACM.

[92] Y. Lin, J.-W. Ahn, P. Brusilovsky, D. He, and W. Real. Imagesieve: Exploratory search of museum archives with named entity-based faceted browsing. *Proceedings of the American Society for Information Science and Technology*, 47(1):1–10, Nov. 2010.

[93] S. Lohmann, J. Ziegler, and L. Tetzlaff. Comparison of tag cloud layouts: Task-related performance and visual exploration. In *INTERACT (1)*, pages 392–404, 2009.

[94] A. Maedche and S. Staab. Measuring similarity between ontologies. In *Proc. Of the European Conference on Knowledge Acquisition and Management - EKAW-2002. Madrid, Spain, October 1-4, 2002*, volume 2473 of *LNCS/LNAI*, Heidelberg, 2002. Springer.

[95] G. Marchionini. Exploratory search: from finding to understanding. *Commun. ACM*, 49:41–46, April 2006.

[96] C. Marlow, M. Naaman, D. Boyd, and M. Davis. Ht06, tagging paper, taxonomy, flickr, academic article, to read. In *Proc. Seventeenth Conference on Hypertext and Hypermedia (Hypertext 2006)*, HT'06, pages 31–40, USA, NY, 2006. ACM.

[97] A. Mehta, A. Saberi, U. Vazirani, and V. Vazirani. Adwords and generalized online matching. *J. ACM*, 54, October 2007.

[98] C. S. Mesnage and M. J. Carman. Tag navigation. In *SoSEA '09: Proceedings of the 2nd international workshop on Social software engineering and applications*, pages 29–32, New York, NY, USA, 2009. ACM.

[99] P. Mika. Ontologies are us: A unified model of social networks and semantics. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(1):5–15, 2007.

[100] S. Milgram. The small world problem. *Psychology Today*, 1:60–67, 1967.

[101] D. Millen and J. Feinberg. Using social tagging to improve social navigation. In *Workshop on the Social Navigation and Community Based Adaptation Technologies. Dublin, Ireland.* Citeseer, 2006.

[102] L. Muchnik, R. Itzhack, S. Solomon, and Y. Louzoun. Self-emergence of knowledge trees: Extraction of the wikipedia hierarchies. *Phys. Rev. E*, 76:016106, Jul 2007.

[103] N. Neubauer and K. Obermayer. Hyperincident connected components of tagging networks. In *HT '09: Proceedings of the 20th ACM conference on Hypertext and hypermedia*, pages 229–238, New York, NY, USA, 2009. ACM.

[104] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[105] O. Nov and C. Ye. Why do people tag?: motivations for photo tagging. *Commun. ACM*, 53:128–131, July 2010.

[106] J. Oomen, L. B. Baltussen, Sander Limonard, M. Brinkerink, A. van Ees, L. Aroyo, J. Vervaart, K. Afsar, and Riste Gligoro. Emerging Practices in the Cultural Heritage Domain - Social Tagging of Audiovisual Heritage. In *Web Science10:Extending the Frontiers of Society On-Line*, Raleigh, NC, USA., 2010.

[107] S. Overell, B. Sigurbjörnsson, and R. van Zwol. Classifying tags using open content resources. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, WSDM '09, pages 64–73, New York, NY, USA, 2009. ACM.

[108] P. Pirolli. An elementary social information foraging model. In *Proceedings of the 27th international conference on Human factors in computing systems*, pages 605–614. ACM, 2009.

[109] A. Plangprasopchok and K. Lerman. Constructing folksonomies from user-specified relations on flickr. In *Proc. of 18th International World Wide Web Conference, WWW '09*, New York, NY, USA, May 2009. ACM.

[110] A. Plangprasopchok, K. Lerman, and L. Getoor. From saplings to a tree: Integrating structured metadata via relational affinity propagation. In *Proceedings of the AAAI workshop on Statistical Relational AI*, Menlo Park, CA, USA, July 2010. AAAI.

[111] M. Ramezani, J. Sandvig, T. Schimoler, J. Gemmell, B. Mobasher, and R. Burke. Evaluating the impact of attacks in collaborative tagging environments. In *Computational Science and Engineering, 2009. CSE '09. International Conference on*, volume 4, pages 136 –143, aug. 2009.

[112] M. Recker and D. Wiley. A non-authoritative educational metadata ontology for filtering and recommending learning objects. *Journal of Interactive Learning Environments*, 9(3):255–271, 2003.

[113] A. Rivadeneira, D. Gruen, M. Muller, and D. Millen. Getting our head in the clouds: toward evaluation studies of tagclouds. *Proceedings of the SIGCHI conference on Human factors in computing systems*, page 998, 2007.

[114] A. W. Rivadeneira, D. M. Gruen, M. J. Muller, and D. R. Millen. Getting our head in the clouds: toward evaluation studies of tagclouds. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 995–998, New York, NY, USA, 2007. ACM.

[115] C. Schmitz, A. Hotho, R. Jäschke, and G. Stumme. Mining association rules in folksonomies. In *Data Science and Classification: Proc. of the 10th IFCS Conf., Studies in Classification, Data Analysis, and Knowledge Organization*, pages 261–270. Springer, 2006.

[116] P. Schmitz. Inducing ontology from flickr tags. In *Proceedings of the Workshop on Collaborative Tagging at WWW2006*, Edinburgh, Scotland, May 2006.

[117] C. Seifert, B. Kump, W. Kienreich, G. Granitzer, and M. Granitzer. On the beauty and usability of tag clouds. In *Information Visualisation, 2008. IV '08. 12th International Conference*, pages 17 –25, july 2008.

[118] S. Sen, S. K. Lam, A. M. Rashid, D. Cosley, D. Frankowski, J. Osterhouse, F. M. Harper, and J. Riedl. tagging, communities, vocabulary, evolution. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work CSCW 06*, volume 4 of *CSCW '06*, pages 181–190. ACM, ACM Press, 2006.

[119] M. A. Serrano, D. Krioukov, and M. Boguñá. Self-similarity of complex networks and hidden metric spaces. *Phys. Rev. Lett.*, 100(7):078701, Feb 2008.

[120] K. Shen and L. Wu. Folksonomy as a complex network, 2005. cite arxiv:cs/0509072.

[121] B. Sigurbjörnsson and R. van Zwol. Tagexplorer: Faceted browsing of flickr photos. In *Technical Report, Yahoo! Research*, 2010.

[122] J. Sinclair and M. Cardew-Hall. The folksonomy tag cloud: when is it useful? *Journal of Information Science*, 34:15, 2008.

[123] A. Singla, R. White, and J. Huang. Studying trailfinding algorithms for enhanced web search. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '10, pages 443–450, New York, NY, USA, 2010. ACM.

[124] G. Solskinnsbakk and J. Gulla. A hybrid approach to constructing tag hierarchies. In *On the Move to Meaningful Internet Systems, OTM 2010*, volume 6427 of *Lecture Notes in Computer Science*, pages 975–982. Springer Berlin / Heidelberg, 2010.

[125] M. Strohmaier. Purpose tagging - capturing user intent to assist goal-oriented social search. In *SSM'08 Workshop on Search in Social Media, in conjunction with CIKM'08*, USA, Napa Valley, 2008.

[126] M. Strohmaier, D. Helic, D. Benz, C. Körner, and R. Kern. Evaluation of folksonomy induction algorithms. *ACM Trans. Intell. Syst. Technol.*, 2012.

[127] M. Strohmaier, C. Koerner, and R. Kern. Why do users tag? Detecting users' motivation for tagging in social tagging systems. In *International AAAI Conference on Weblogs and Social Media (ICWSM2010), Washington, DC, USA, May 23-26*, 2010.

[128] M. Strohmaier, C. Trattner, and D. Helic. The benefits and limitations of tag clouds as a tool for social navigation from a network-theoretic perspective. *Journal of Universal Computer Science*, 2010. (UNDER REVIEW).

[129] C. Trattner. Querycloud: Automatically linking related documents via search query (tag) clouds. In *Proceedings of IADIS International Conference WWW/Internet 2010*, ICWI'10. IADIS, 2010.

[130] C. Trattner. Improving tag cloud navigability: A comparative study. In *Proc. of the 33th International Conference on Information Technology Interfaces*, Cavtat, Dubrovnik, Croatia, 2011. IEEE.

[131] C. Trattner. Improving the navigability of tagging systems with hierarchically constructed resource lists and tag trails. *CIT*, 19(3):155–167, 2011.

[132] C. Trattner. Linking related content in web encyclopedias with search query tag clouds. *IJWI*, 9(1):33–55, 2011.

[133] C. Trattner, H. Denis, and M. Strohmaier. On the construction of efficiently navigable tag clouds using knowledge from structured web content. *Journal of Universal Computer Science*, 2011.

[134] C. Trattner, I. Hasani-Mavriqi, D. Helic, and H. Leitner. The austrian way of wiki(pedia)!: development of a structured wiki-based encyclopedia within a local austrian context. In *Proceedings of the 6th International Symposium on Wikis and Open Collaboration*, WikiSym '10, pages 1–10, New York, NY, USA, 2010. ACM.

[135] C. Trattner and D. Helic. Extending the basic tagging model: Context aware tagging. In *IADIS International Conference WWWInternet 2009*, pages 76 – 83. IADIS, 2009.

[136] C. Trattner and D. Helic. Linking related documents: combining tag clouds and search queries. In *Proceedings of the 10th international conference on Web engineering*, ICWE'10, pages 486–489, Berlin, Heidelberg, 2010. Springer-Verlag.

[137] C. Trattner, D. Helic, and S. Maglajlic. Enriching tagging systems with google query tags. In *Information Technology Interfaces (ITI), 2010 32nd International Conference on*, pages 205 –210, june 2010.

[138] C. Trattner, C. Körner, and D. Helic. Enhancing the navigability of social tagging systems with tag taxonomies. In *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, i-KNOW '11, pages 18:1–18:8, New York, NY, USA, 2011. ACM.

[139] C. Trattner, Y.-L. Lin, D. Parra, W. Real, and P. Brusilovsky. Evaluating tag-based information access in image collections. In *Proceedings of the 23rd ACM Conference on Hypertext and Social Media*, HT '12, New York, NY, USA, 2012. ACM.

[140] C. Trattner, P. Singer, D. Helic, and M. Strohmaier. Exploring the differences and similarities between hierarchical decentralized search and human navigation in infromation networks. In *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, Iknow '12, New York, NY, USA, 2012. ACM.

[141] C. Trattner, M. Strohmaier, D. Helic, and K. Andrews. The benefits and limitations of tag clouds as a tool for social navigation. *Technical Report – IICM, Graz University of Technology*, 2011.

[142] P. Vakkari, M. Pennanen, and S. Serola. Changes of search terms and tactics while writing a research proposal a longitudinal case study. *Inf. Process. Manage.*, 39:445–463, May 2003.

[143] P. Venetis, G. Koutrika, and H. Garcia-Molina. On the selection of tags for tag clouds. In *Proceedings of the fourth ACM international conference on Web search and data mining*, WSDM '11, pages 835–844, New York, NY, USA, 2011. ACM.

[144] C. Wartena. Automatic classification of social tags. In *Proceedings of the 14th European conference on Research and advanced technology for digital libraries*, ECDL'10, pages 176–183, Berlin, Heidelberg, 2010. Springer-Verlag.

[145] D. J. Watts, P. S. Dodds, and M. E. J. Newman. Identity and search in social networks. *Science*, 296:1302–1305, 2002.

[146] D. J. Watts and S. H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393(6684):440–442, June 1998.

[147] R. West and J. Leskovec. Automatic versus human navigation in information networks. In *Proceedings of the 6th International AAAI Conference on Weblogs and Social Media*, ICWSM'12, Dublin, Ireland, 2012. AAAI.

[148] R. West and J. Leskovec. Human wayfinding in information networks. In *Proceedings of the 21th international conference companion on World wide web*, WWW'12, New York, NY, USA, 2012. ACM.

[149] R. W. White and S. M. Drucker. Investigating behavioral variability in web search. In *Proceedings of the 16th international conference on World Wide Web*, WWW '07, pages 21–30, New York, NY, USA, 2007. ACM.

[150] Wikipedia. Wikipedia – Delicious. http://en.wikipedia.org/wiki/Delicious_(website), 2012. [Online; accessed 2012-05-21].

[151] X. Wu, L. Zhang, and Y. Yu. Exploring social annotations for the semantic web. *Proceedings of the 15th international conference on World Wide Web*, page 426, 2006.

[152] Y. Zhang, B. Jansen, and A. Spink. Time series analysis of a Web search engine transaction log. *Information Processing & Management*, 45(2):230–245, 2009.

[153] A. Zubiaga. Enhancing Navigation on Wikipedia with Social Tags. *Wikimania 2009*, pages 1–9, 2009.

[154] A. Zubiaga, C. Körner, and M. Strohmaier. Tags vs shelves: from social tagging to social classification. In *Proceedings of the 22nd ACM conference on Hypertext and hypermedia*, HT '11, pages 93–102, New York, NY, USA, 2011. ACM.

A

Exploring Differences and Similarities of Hierarchical Decentralized Search and Human Navigation

This chapter is based on the paper *"Exploring Differences and Similarities between Hierarchical Decentralized Search and Human Navigation in Information Networks"* which was presented at the 12th International Conference on Knowledge Management and Knowledge Technologies in 2012.

This paper presents a large scale study exploring the differences and similarities of hierarchical decentralizes search and human navigation in information networks based on the entire English Wikipedia graph. The intention of including this paper in this dissertation is to shed light on the question to which extent it is justified to simulate human navigational behavior in information networks (such as tagging systems) with a hierarchical decentralized search procedure.

The original contribution was published in the proceedings of the conference and can be found in [26].

A.1 Abstract

Decentralized search in networks is an activity that is often performed in online tasks. It refers to situations where a user has no global knowledge of a network's topology, but only local knowledge. In Wikipedia for instance, humans typically have local knowledge of the links emanating from a given Wikipedia article, but no global knowledge of the entire Wikipedia graph. This makes the task of navigation to a target Wikipedia article from a given starting article an interesting problem for both humans and algorithms. As we know from previous studies, people can have very efficient decentralized search procedures that find shortest paths in many cases, using intuition

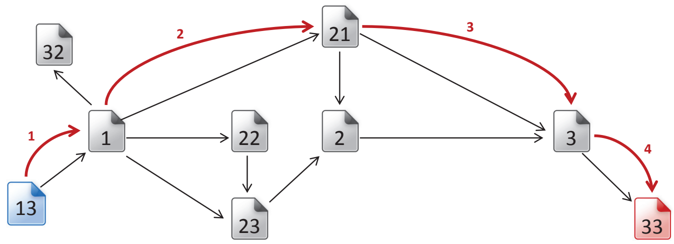
about a given network. This intuition can be modeled as hierarchical background knowledge that people access to approximate a networks' topology. In this paper, we explore the differences and similarities between decentralized search that utilizes hierarchical background knowledge and actual human navigation in information networks. For that purpose we perform a large scale study on the Wikipedia information network with over 500,000 users and 1,500,000 click trails. As our results reveal, a decentralized search procedure based on hierarchies created directly from the link structure of the information network simulates human navigational behavior better than simulations based on hierarchies that are created from external knowledge.

A.2 Introduction

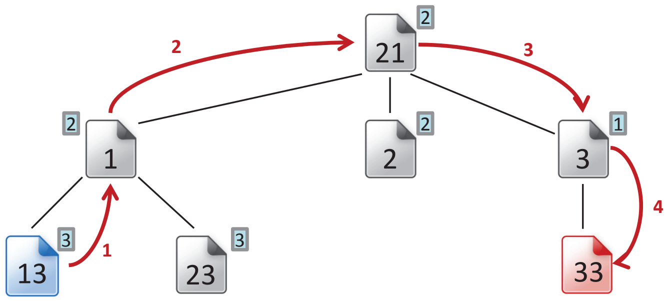
In 1967, Milgram conducted his now famous small-world experiment [18], in which randomly selected people from Nebraska had to pass on a letter to a specific target person in Boston. The specific experimental setup required the participants to pass the letter in a decentralized manner, i.e. they were only allowed to pass the letter through their local social networks. Despite this restriction, the average chain length of those letters that reached the target person was only six - thus, giving rise to the hypothesis that the USA constituted a small-world.

One of the most interesting research questions raised by this experiment was to understand and characterize the algorithm that people use to efficiently find other distant people in social networks. To that end, among others, Kleinberg introduced the theory of *decentralized search* and provided a theoretical explanation of this human ability [14, 15, 16]. In a number of studies Kleinberg showed that social networks possess certain latent structural properties that humans are aware of and are able to utilize in their search for other people. This allows them to find short paths between two arbitrary network nodes efficiently even with only local knowledge of the network. Consequently, Kleinberg also examined the structure of such latent structural properties that he called background knowledge, and discovered that social networks can be efficiently searched, i.e. in $log(N)$, where N are the number of nodes in the network, if the nodes of the network can be organized into a hierarchy. This theoretical model is also known as Kleinberg's hierarchical network model [16].

Based on these ideas, Lada Admic [1] implemented a decentralized search algorithm that utilizes hierarchical background knowledge of a network and applied that algorithm in a number of experiments. Adamic showed that the algorithm performs well in simulating human-like search behavior in social networks. Furthermore, she demonstrated that the performance of the simulator depends on the quality of the background knowledge of the network.



(a) Information network



(b) Hierarchical background knowledge

Figure A.1: An example of decentralized search in an information network (a) using hierarchical background knowledge of this network (b). The information network links information for instance document pages (i.e., Wikipedia pages) with each other. The search begins at the blue node 13. The destination node is the red node 33. At each step, the search algorithm selects one of the current node's adjacent nodes, which is the closest to the target node in the hierarchy. The numbers in boxes in (b) provide the distance between the current node and the destination node 33. At step one, node 13 has a single adjacent node 1, so search continues to 1. At step two, 1's adjacent nodes include 21, 22, 23 and 32. The algorithm consults the hierarchy finding out that node 21 is the closest to the destination node. At step three, the algorithm has an option to move to nodes 2 or 3. The simulation selects node 3, since again, it has the smallest distance to the destination node. Finally, at step 4, the target node is successfully reached.

In our previous work [11, 9, 10], we applied a variant of Adamic's algorithm for simulation of navigation in information networks. Navigation in information networks is a kind of decentralized search, as users at each particular step of their navigation are only aware of links emanating from the current document. Thus, this situation is intuitively very similar to decentralized search in social networks. For example, in [11] we developed a hierarchical decentralized search algorithm based on the ideas of Adamic that allows decentralized search in social tagging systems. By constructing tag hierarchies from the bipartite tag-resource network structures of a number of tagging systems and by using this background knowledge as input for our hi-

Algorithm 12 Hierarchical decentralized searcher

```
INPUT: network N, hierarchy H, start-node s, target-node t
c ← s
while c ≠ t do
  o ← −1
  dist_min ← ∞
  /* Γ(c) is a set of all neighbors of c */
  for each n ∈ Γ(c) do
    dist ← h(n, H) + h(w, H) − 2h(n, w, H) − 1
    if dist < dist_min then
      dist_min ← dist
      o ← n
    end if
  end for
  c ← o
end while
```

erarchical decentralized search algorithm, we could show that tag hierarchies perform extremely well in searching social tagging systems. In subsequent work [22], we also demonstrated that the most semantically sound tag hierarchies are also those that perform well on navigational tasks. However, our previous experiments were based on intuition how humans navigate and we have not yet compared our simulations (based on decentralized search) with real human navigation paths.

Hence, the purpose of this paper is to compare simulations based on hierarchical decentralized search with a large-scale corpus of human navigational paths and to reveal whether or not it is justified to simulate human navigational behavior in information networks with the hierarchical decentralized search procedure as introduced and used by us in previous work [12, 22, 23, 24]. To that end, we compared more than 150,000 click trails of users navigating the complete English Wikipedia with simulations. As our results reveal, decentralized search procedures based on hierarchies created directly from the link structure of the information network simulate human navigational behavior better than simulations based on hierarchies that are created from external knowledge.

The remainder of the paper is structured as follows: In Section A.3, we discuss related work. In Section A.4 we shortly present our simulation model for user navigation in information networks. In Section A.5, we outline our experimental setup and in Section A.6 we present the results. Finally, Section A.7 concludes the paper.

A.3 Related Work

Related work in this area can be broadly divided into the following three areas: Web click-trail analysis, navigation in complex networks and hierarchy creation from networks.

A.3.1 Click-Trail Analysis

Click-trail analysis has been mainly performed to improve the Web search results of users. For instance, in [5, 21] the authors assessed the possibility to rank search results more efficiently by taking the users click-trails into account. In [2] a large scale study was conducted to investigate how often users revisit the same Web page. To the best of our knowledge, there is only one study that tries to understand how people navigate in information networks by analyzing a large click-trail log from the online game Wikispeedia[1]. In [28] West and Leskovec performed a study of users navigating Wikipedia articles[2]. In their work they found out that user navigation behavior is close to the short paths of the network. In subsequent work [27], the authors analyzed a number of decentralized search algorithms and benchmarked them against their human click corpus. The most interesting result was that even simple search strategies such as utilizing node degrees, outperform human information seeking. Contrary to the work of West and Leskovec, our study is not focused on finding the fastest decentralized search strategy based on machine learning algorithms, instead we are interested to investigate to what extent it is justified to simulate human navigation in information networks with hierarchical decentralizes search.

A.3.2 Navigation in Networks

Research on navigation in complex networks was initiated by the famous small-world experiment conducted by Milgram [18]. Apart from the work on the algorithmic perspective of search in social networks that we mention in Section A.2, a number of studies recently dealt with navigability of other types of complex networks. In [20], the authors extend the notion of Kleinberg's background knowledge to the notion of *hidden metric spaces*. In such hidden metric spaces nodes are identified by their co-ordinates – distance between nodes is their geometric distance in a particular metric space. Navigation strategies in complex networks are then based on the distances between nodes – an agent always navigates to the node with the smallest distance to a particular destination node. An interesting research question is the structure of such hidden metric spaces that underlie observable networks. In [6], the authors introduce a model with the circle as a hidden

[1] http://www.cs.mcgill.ca/~rwest/wikispeedia/
[2] http://schools-wikipedia.org/

metric space and show its effects on routing in the global airport network. In [17] the authors discuss hyperbolic geometry as a hidden metric space (which can be approximated by a node hierarchy), whereas in [7] the authors apply hyperbolic geometry as a model of the hidden metric space of the Internet and design a novel greedy Internet routing algorithm. In this work we will focus on Kleinberg's hierarchical network model.

A.3.3 Extracting Hierarchies from Networks

Hierarchies that are extracted from networks play an important role in many of these network navigation models. Apart from the tag hierarchy induction algorithms based on bipartite networks such as e.g., [13, 3, 10], researchers also proposed hierarchy extraction algorithms for general networks. In [19] the authors discuss an algorithm for hierarchy construction in Wikipedia networks based on metrics for estimating hierarchy level of single nodes. Also, Clauset et al. [8] present a hierarchy induction algorithm based on prediction of hierarchical links. To extract hierarchical background knowledge as hidden metric space for our decentralized search algorithm, we rely on the hierarchy induction algorithms of [13, 19] in this paper.

A.4 The Algorithm

To simulate human information seeking behavior in information networks, we implemented in the past a hierarchical search algorithm (see Algorithm 12) based on the ideas of Lada Adamic. The algorithm takes as input a given network, start and target nodes and a hierarchical representation of the given network. To navigate from one node in the network to another, all adjacent nodes of the current node are examined and the distance to the target node is calculated over the input hierarchy. The simulator then selects as the next step the node with the minimal distance to the target which is calculated over the given input hierarchy (see Figure A.1). Please note that the pseudo code of our algorithm does not include the cancellation strategy. This is done if the simulator re-visits a node. However, as shown by [28] only a small fraction of users choose the same link again for navigating from one resource to another in an information network. For that purpose, we ignore back tracking. We also cancel search in the case we cannot find a particular node of the network in hierarchy. When the distance function returns the same minimum distance for more than one adjacent node, we try to avoid the nodes that we already visited. To simplify the pseudo code in Algorithm 12, we omit this avoiding strategy from the code.

A.5 Experimental Setup

The following section discusses in detail the experimental design used to evaluate our approach of hierarchical decentralized search to simulate human navigational behavior in information networks.

A.5.1 Datasets

Wikipedia Click Dataset

In order to compare the behavior of the search algorithm with human navigation, we analyze a click dataset from the complete English Wikipedia. The dataset comes from the online platform the Wikigame[3]. There are two reasons for our decision to use this kind of dataset. First, there are no freely available datasets that include complete click paths from a specific start node to a specific target node. Typically, one has to apply heuristics to extract users, their sessions, and their click trails. In Wikigame, we have a complete sequence of clicks of different users participating in a game that requires from the users to navigate from e.g., *"Wolfgang Amadeus Mozart"* to e.g., *"Arnold Schwarzenegger"*. In turn, other datasets do not include explicit $(start, target)$ information. The second reason is basically the large scale of the dataset, with records of more than 500,000 users and 1,500,000 click trails. However, for the purposes of this study we analyze only a subset of this large scale dataset.

Wikipedia Network Dataset

Additionally to the dataset record of Wikigame click paths, our work is based on an information network dataset (= directed link-network dataset) of the English-Wikipedia from February 2012. We use this kind of dataset as the basis for our simulations. All in all, the dataset includes around 10,000,000 articles and around 250,000,000 links.

Wikipedia Category Label Datasets

Since our decentralized search simulations are based on hierarchical background knowledge of the information network, the question arises how can we extract this kind of knowledge from our Wikipedia dataset. A simple idea is to use Wikipedia category labels for constructing a hierarchy representation of the network. Another idea is to use external meta-data information, such as social tags which were shown useful to classify information such as Web pages [29, 30]. In our case, we used a dataset of Wikipedia category labels as well as a dataset of social tags from Delicious which only

[3] http://thewikigame.com/

consists of annotated Wikipedia articles. Overall, the Wikipedia category label dataset includes around 2,300,000 category labels, 4,500,000 articles and 30,000,000 category label assignments. The Delicious tag dataset includes around 440,000 tags, 580,000 articles and 3,400,000 tag assignments.

A.5.2 Preliminaries

Click-Trail Selection

For the purpose of our study, we only considered games (=click trails) that were successfully accomplished. We also selected only those click trails where the start and target node were present in all of the hierarchies that we produced. At the end, we analyzed over 150,000 click trails.

Creating Hierarchies

In previous work [22] we showed that our algorithm depends on the quality of the hierarchical knowledge extracted from the information network. As also shown, the best results are archived by creating hierarchies that are created by graph based clustering algorithms that are based on the tag network's tag co-occurrence graph. In this work, we use two different types of hierarchy induction algorithms. One based on the ideas expressed earlier and one new algorithm that only considers in- and out-degree of the nodes of the information network.

Creating Hierarchies from External Knowledge: The first approach we use is based on the ideas of [13]. In their work the authors introduce a generic algorithm for producing hierarchies from bipartite networks such as tag-to-resource networks. The algorithm can be applied to arbitrary bipartite structures. The algorithm takes two parameters as input. The first is a ranked list of tags sorted by their centrality in the projected tag-to-tag network. This centrality ranking acts as a proxy to the generality ranking of tags. Benz et al. [4] showed that the centrality provides a viable approximation for term abstractness in tags. The second input parameter is the tag similarity matrix. The algorithm starts then by a single node hierarchy with the most general tag as the root node and then iterates through the centrality list. At each iteration step, the algorithm adds the current tag to the hierarchy as a child to its most similar tag. The centrality and similarity measure are exchangeable – in [13] the authors use closeness centrality and cosine similarity, whereas in [3] the authors select degree centrality and co-occurrence similarity measure. As both combinations perform similarly in supporting navigation [12], we select in this work the latter combination because of better computational properties. Furthermore, we adopted the algorithm of Benz et al. to produce a resource taxonomy instead of a tag taxonomy. We achieve this by simply switching our computations from the projected tag-to-tag network to the projected resource-to-resource network. This algorithm is then applied to generate a Wikipedia resource hierarchy on the basis of the Delicious tag dataset as well on the basis of the Wikipedia category label dataset.

Creating Hierarchies from the Network: The second type of hierarchy we produce for our simulation is based on the ideas of [19]. The algorithm is based on the idea that each network possesses an inherent hierarchical structure that leads to the emergence of observable structural properties such as power-law degree distributions and high node clustering (cf. [8]). The algorithm then aims to recognize and extract that hierarchical structure. Thus, the algorithm iterates through all links in the network and decides – using a simple criteria – if that link is of a hierarchical type, in which case it remains in the network, or if that link is of some other kind (e.g., a synonym link), in which case the link is removed from the network. To that end, the algorithm assigns to each node a so-called hierarchical score, which is a measure stating the generality of a node. For each link the ratio between hierarchical scores of two incident to that link is calculated. The simple idea is that if that ratio is close to one then those two nodes are very close in their generality and they are situated in the same hierarchy level – thus, the link between those two nodes is not a hierarchical one and is therefore removed from the network. Similarly, if the hierarchical ratio for a link is close to zero then those two nodes are very far away from each other in the hierarchy and the link is removed (e.g., an article in a very small town in the USA, say Paris, Texas, links to the article in the United States). Technically, the authors define two thresholds – high and low thresholds – to decide on the removal of the links. Thus, a link is removed if the hierarchical ratio is greater than the high threshold or smaller than the low threshold. Another technical issues is the decision on how to calculate the hierarchical score. In their paper, the authors compare five different hierarchical scores ranging from global scores such as betweenness centrality to local scores such as ratio of in-degree and out-degree of a node. In our experiments we use a local score, defined as:

$$hs(n) = \frac{d_{in}(n)}{d_{out}(n)} \sqrt{d_{in}(n)}. \tag{A.1}$$

The term $\sqrt{d_{in}(n)}$ ensures that a node having e.g., 200 in-degree and 100 out-degree is rendered more general than a node having e.g., 2 in-degree and 1 out-degree. As thresholds we choose 0.6 and 0.2 for high and low thresholds respectively (cf. [19]).

A.5.3 Measures

To compare our simulations with human navigation, we define a number of measures. In the following list, we give a short overview of these measures and how they are calculated:

- **Success Rate:** As discussed before, we use in our analysis only successful games (=click trails), i.e., the success rate of human navigators is 100%. Since we perform our simulations on the same search trails, we can identify with this measure to which extent the simulation differs from reaching the destination node in each step or on average. In our analysis we calculate the mean local s and global (=overall) success rate s_g.

- **Number of Hops:** Another interesting measure is the number of hops needed to reach the target node. We capture this on a global basis $\overline{h}$.

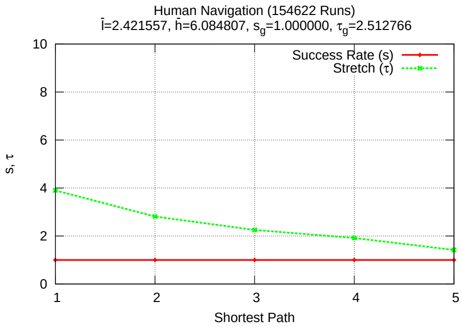
- **Stretch:** Stretch captures the ratio of the number of steps and the global shortest path. As shown in [28] humans are typically very efficient at finding shortest paths. On average, they find information in Wikipedia in not more than two more steps than the shortest possible path. Thus, with this measure we identify how good our simulation is in finding shortest paths in each step τ and on average overall τ_g compared to human navigators.

- **Path Similarity:** We calculate path similarity to determine the extent to which successful paths of our simulations differ from real user navigational trails. Since the user's click paths in general show a high diversity by terms of similarity (see Figure A.3(a)), we calculate path similarity as
$$\frac{ctr(h)_{(a,b)} \cap ctr(s)_{(a,b)}}{ctr(s)_{(a,b)}} \tag{A.2}$$
where $ctr(h)_{(a,b)}$ is the set of human click trails for the search pair (a, b) and where $ctr(s)_{(a,b)}$ is the set of simulation trails for the same pair.

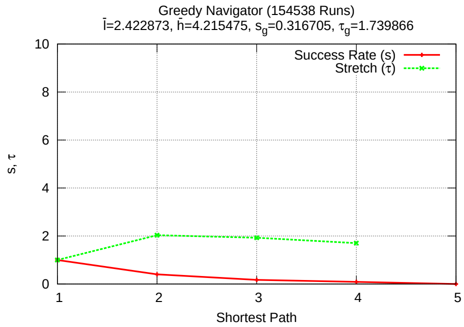
- **Degree:** Finally, we also investigate the median in- and out-degree values of the nodes visited by the simulator and the human navigator (we use the median in this case since the values are not normally distributed).

A.6 Results

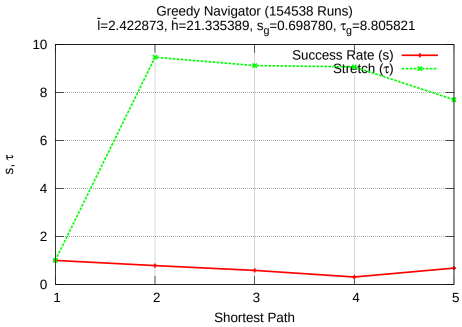
We simulated over 150,000 searches on the Wikipedia link-network utilizing three different hierarchies as background knowledge for our hierarchical decentralized search procedure. To make results comparable, we run our



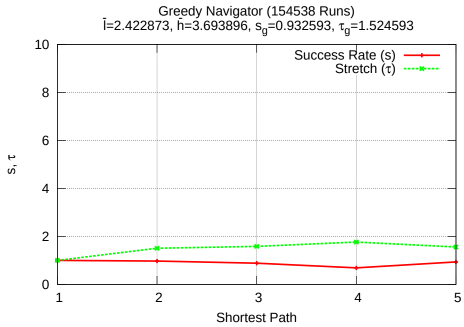
(a) Humans navigators



(b) Searcher with Wikipedia category hierarchy



(c) Searcher with Delicious tag hierarchy



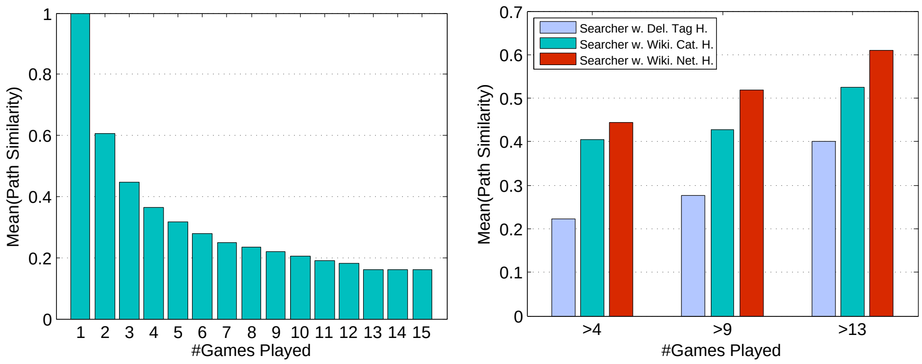
(d) Searcher with Wikipedia network hierarchy

Figure A.2: Results of Human navigators vs. simulator (=greedy navigator): Average shortest path $\bar{l}$, average hop count $\bar{h}$, success rate s and stretch τ – global average values (s_g and τ_g) and distribution over shortest paths. As shown, the simulator with the hierarchy based on the Wikipedia link structure simulates humans best (highest success rate $s_g = 0.93$ and stretch that is close to the human navigators. The simulator with the Wikipedia category label hierarchy performs worst, success rate is only $s_g = 0.31$.

simulations on the Wikipedia link-network using only the start target node pairs as present in the human click-trail dataset.

A.6.1 Success Rate, Number of Hops and Stretch

In Figure A.2 we illustrate the first results of our comparative evaluation. As shown, the simulator utilizing the hierarchy based on the Wikipedia link structure generates the best results. We can observe the highest success rate $s_g = 0.93$ of all other simulators. The worst performance $s_g = 0.31$ is achieved by the simulator with hierarchical background knowledge generated from the Wikipedia category labels. Interestingly, the success rate of



(a) Path similarity between human navigators

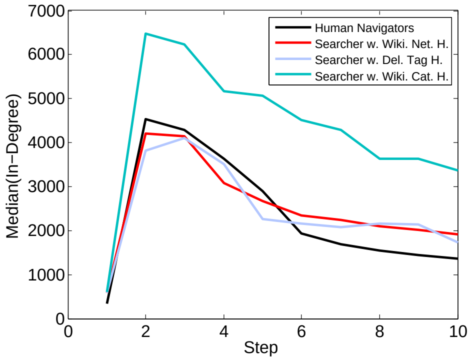
(b) Path similarity between humans and the simulator

Figure A.3: Path Similarity between human navigators (a) and path similarity between humans and the simulator (b). As shown, in Figure (a) path similarity drops significantly the more people play the same game. For games that are played more than 13 times, the path similarity drops down to 18% (but also seems to stay steady) indicating that humans agree little on taking the same paths to reach the target node. In Figure (b), path similarity between humans and the simulator with different background knowledge is shown. We can observe that the searcher utilizing the Wikipedia network hierarchy as background knowledge simulates human navigational behavior best.

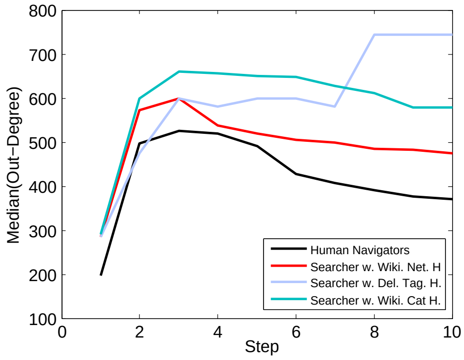
the simulations based on the Delicious tag hierarchy is quite high, taking into account that the Delicious tag dataset covers five times less articles in Wikipedia. This leads to the situation that the Delicious hierarchy contains also five times fewer nodes than the hierarchy extracted from the Wikipedia category labels, which means that the simulation is more likely to fail the search, since a possible selected node of the simulation is not present in the hierarchy. However, as also shown in Figure A.2 the average hop length is high $\overline{h} = 21.34$. This demonstrates that it is possible to navigate successfully through an information network even if the hierarchy is not complete. On the other hand, we can see that hierarchies directly extracted from the information network are better suited as hierarchical background knowledge than hierarchies based on external knowledge.

A.6.2 Path Similarity

In addition to the previous results, we illustrate in Figure A.3(a) path similarity between the human navigators. As shown, the more games are played the more diverse the paths of the users are, i.e., humans have only little agreement on how they route through an information network. This could be explained by their familiarity regarding the search item or their experi-



(a) Median in-degree distribution



(b) Median out-degree distribution

Figure A.4: Median in and out-degree distributions for human navigators and simulations. As shown again, the hierarchical decentralized searcher utilizing the Wikipedia network hierarchy as background knowledge simulates human search behavior best. However, as shown in Figure (b), the simulator in general favors higher out-degree nodes than human navigators.

ence with the system [25]. In Figure A.3(b), we compare the similarity of the successful paths conducted by human navigators and the ones resulting from our simulator on different hierarchies. As the results reveal, again simulations based on the Wikipedia network hierarchy are most similar to human navigational paths.

A.6.3 Degree

Finally, Figure A.4 shows the median in- and out-degree distributions for human navigators and simulations. As observed in related work by West and Leskovec [28], humans follow certain patterns in their information seeking behavior. In particular, high degree nodes are typically used in the first steps of the search, while similar nodes are used by the end of the search. Since degree is highly correlated to similarity [28], we only focus on degree in our analysis. As shown in Figure A.4, humans as well as simulators choose high degree nodes in the first step of their search, while they tend to utilize low-degree nodes at the end of the search procedure. Again we can see that the hierarchical decentralized simulator utilizing the Wikipedia network hierarchy as background knowledge is most similar to human search behavior. Simulations based on the Wikipedia category label hierarchy perform worst in this case. This behavior might be an explanation for the bad performance of this searcher in terms of success rate and stretch as shown in Figure A.2.

A.7 Conclusions and Outlook

In this work we explored the differences and similarities between hierarchical decentralized search and human navigational behavior in information networks and to reveal whether or not it is justified to simulate human navigational behavior in information networks with the hierarchical decentralized search procedure introduced and used by us in previous work [12, 22, 23, 24]. Based on a large-scale click dataset of over 150,000 click trails from the online platform the Wikigame, we performed a number of experiments to gain insights into how humans search in information networks and how well simulations based on hierarchical decentralized search correlate with humans click trails. Generating background knowledge from various sources, we could show that a decentralized search procedure based on hierarchies created directly from the link structure of the information network simulates human navigational behavior better than simulations based on hierarchies that are created from external knowledge.

Limitations and Future Work: Even if the paper presents a large-scale study on how to simulate human navigational behavior in information networks with decentralized search, we have to acknowledge that research in this context is still at an early stage and therefore also has some limitations. One of these limitations which we would like to study in the near future work is the fact that the present study does not include a comparison of hierarchical decentralized search with other well-known probabilistic approaches such as *markov-chains*. However, even if our presented method showed good results in simulating human navigational behavior it would be interesting to see whether or not other approaches perform better in simulating human information seeking behavior of humans in information networks than our current algorithm.

References

[1] L. Adamic and E. Adar. How to search a social network. *Social Networks*, 27(3):187 – 203, 2005.

[2] E. Adar, J. Teevan, and S. T. Dumais. Large scale analysis of web revisitation patterns. In *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, CHI '08, pages 1197–1206, New York, NY, USA, 2008. ACM.

[3] D. Benz, A. Hotho, and G. Stumme. Semantics made by you and me: Self-emerging ontologies can capture the diversity of shared knowledge. In *Proc. of the 2nd Web Science Conference (WebSci10)*, Raleigh, NC, USA, 2010.

[4] D. Benz, C. Körner, A. Hotho, G. Stumme, and M. Strohmaier. One tag to bind them all : Measuring term abstractness in social metadata. In *Proceedings of the 8th Extended Semantic Web Conference (ESWC 2011)*, Heraklion, Crete, May 2011.

[5] M. Bilenko and R. W. White. Mining the search trails of surfing crowds: identifying relevant websites from user activity. In *Proceedings of the 17th international conference on World Wide Web*, WWW '08, pages 51–60, New York, NY, USA, 2008. ACM.

[6] M. Boguñá, D. Krioukov, and K. C. Claffy. Navigability of complex networks. *Nature Physics*, 5:74–80, Jan. 2009.

[7] M. Boguñá, F. Papadopoulos, and D. Krioukov. Sustaining the Internet with hyperbolic mapping. *Nature Communications*, 1:62, Sept. 2010.

[8] A. Clauset, C. Moore, and M. E. J. Newman. Hierarchical structure and the prediction of missing links in networks. *Nature*, 453(7191):98–101, 2008.

[9] D. Helic, , C. Körner, M. Granitzer, M. Strohmaier, and C. Trattner. Navigational efficiency of broad vs. narrow folksonomies. In *Proceedings of the 23rd ACM Conference on Hypertext and Social Media*, HT '12, New York, NY, USA, 2012. ACM.

[10] D. Helic and M. Strohmaier. Building directories for social tagging systems. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, CIKM '11, pages 525–534, New York, NY, USA, 2011. ACM.

[11] D. Helic, M. Strohmaier, C. Trattner, M. Muhr, and K. Lermann. Pragmatic evaluation of folksonomies. In *Proc. of the 21st International World Wide Web conference, WWW '11*, New York, NY, USA, 2011. ACM.

[12] D. Helic, C. Trattner, M. Strohmaier, and K. Andrews. Are tag clouds useful for navigation? a network-theoretic analysis. *International Journal of Social Computing and Cyber-Physical Systems*, 2011.

[13] P. Heymann and H. Garcia-Molina. Collaborative creation of communal hierarchical taxonomies in social tagging systems. Technical Report 2006-10, Stanford InfoLab, April 2006.

[14] J. Kleinberg. The small-world phenomenon: An algorithmic perspective. In *Proceedings of the 32nd ACM Symposium on Theory of Computing*, 2000.

[15] J. M. Kleinberg. Navigation in a small world. *Nature*, 406(6798), August 2000.

[16] J. M. Kleinberg. Small-world phenomena and the dynamics of information. In *Advances in Neural Information Processing Systems (NIPS) 14*, page 2001. MIT Press, 2001.

[17] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá. Hyperbolic geometry of complex networks. *Phys. Rev. E*, 82(3):036106, Sep 2010.

[18] S. Milgram. The small world problem. *Psychology Today*, 1:60–67, 1967.

[19] L. Muchnik, R. Itzhack, S. Solomon, and Y. Louzoun. Self-emergence of knowledge trees: Extraction of the wikipedia hierarchies. *Phys. Rev. E*, 76:016106, Jul 2007.

[20] M. A. Serrano, D. Krioukov, and M. Boguñá. Self-similarity of complex networks and hidden metric spaces. *Phys. Rev. Lett.*, 100(7):078701, Feb 2008.

[21] A. Singla, R. White, and J. Huang. Studying trailfinding algorithms for enhanced web search. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '10, pages 443–450, New York, NY, USA, 2010. ACM.

[22] M. Strohmaier, D. Helic, D. Benz, C. Körner, and R. Kern. Evaluation of folksonomy induction algorithms. *ACM Trans. Intell. Syst. Technol.*, 2012.

[23] C. Trattner. Improving the navigability of tagging systems with hierarchically constructed resource lists and tag trails. *CIT*, 19(3):155–167, 2011.

[24] C. Trattner, H. Denis, and M. Strohmaier. On the construction of efficiently navigable tag clouds using knowledge from structured web content. *Journal of Universal Computer Science*, 2011.

[25] C. Trattner, Y.-L. Lin, D. Parra, W. Real, and P. Brusilovsky. Evaluating tag-based information access in image collections. In *Proceedings of the 23rd ACM Conference on Hypertext and Social Media*, HT '12, New York, NY, USA, 2012. ACM.

[26] C. Trattner, P. Singer, D. Helic, and M. Strohmaier. Exploring the differences and similarities between hierarchical decentralized search and human navigation in infromation networks. In *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, Iknow '12, New York, NY, USA, 2012. ACM.

[27] R. West and J. Leskovec. Automatic versus human navigation in information networks. In *Proceedings of the 6th International AAAI Conference on Weblogs and Social Media*, ICWSM'12, Dublin, Ireland, 2012. AAAI.

[28] R. West and J. Leskovec. Human wayfinding in information networks. In *Proceedings of the 21th international conference companion on World wide web*, WWW'12, New York, NY, USA, 2012. ACM.

[29] A. Zubiaga. Enhancing Navigation on Wikipedia with Social Tags. *Wikimania 2009*, pages 1–9, 2009.

[30] A. Zubiaga, C. Körner, and M. Strohmaier. Tags vs shelves: from social tagging to social classification. In *Proceedings of the 22nd ACM conference on Hypertext and hypermedia*, HT '11, pages 93–102, New York, NY, USA, 2011. ACM.

B

A Complete List of Own Publications

B.1 Peer-Reviewed Journal Publications

J1. Trattner, C. and Kappe, F. 2012. *Social Stream Marketing on Facebook: A Case Study*. International Journal of Social and Humanistic Computing (IJSHC).

J2. Trattner, C. 2011. *Linking Related Content in Web Encyclopedias with search query tag clouds*. In the International Journal on WWW/Internet, Volume 9, Issue 2, pp. 33-50.

J3. Trattner, C. 2011. *Improving the Navigability of Tagging Systems with Hierarchically Constructed Resource Lists and Tag Trails*. In the Journal of Computing and Information Technology, Volume 19, Issue 3, pp. 155-167.

J4. Trattner, C., Helic, D. and Strohmaier, M. 2011. *On the Construction of Efficiently Navigable Tag Clouds Using Knowledge From Structured Web Content*. In the Journal of Universal Computer Science, Volume 17, Issue 4, pp. 565-582.

J5. Helic, D., Trattner, C., Strohmaier, M. and Andrews, K. 2011. *Are Tag Clouds Useful for Navigation? A Network-Theoretic Analysis*. In the Journal of Social Computing and Cyber-Physical Systems, Volume 1, Issue 1, pp. 33-55.

B.2 Peer-Reviewed Conference Publications

C1. Trattner, C., Lin, Y., Parra, D., Yue, Z., Real, W. and Brusilovsky, P. 2012. *Evaluating Tag-Based Information Access in Image Collections*. In Proceedings of the 23rd ACM Conference on Hypertext and Social Media (HT 2012), ACM, New York, NY, USA, pp. 113-122.

C2. Trattner, C., Singer, P., Helic, D. and Strohmaier, M. 2012. *Exploring Differences and Similarities between Hierarchical Decentralized Search and Human Navigation in Information Networks.* In Proceedings of the 12th International Conference on Knowledge Management and Knowledge Technologies (I-Know 2012), ACM, New York, NY, USA.

C3. Helic, D., Körner, C., Granitzer, M., Strohmaier, M. and Trattner, C. 2012. *Navigational Efficiency of Broad vs. Narrow Folksonomies.* In Proceedings of the 23rd ACM Conference on Hypertext and Social Media (HT 2012), ACM, New York, NY, USA, pp. 63-72.

C4. Trattner, C. and Kappe, F. 2011. *Social Media Marketing with Facebook: A Case Study.* In Proceedings of IADIS International Conference WWW/Internet 2011, IADIS, Rio de Janeiro, Brasil, pp. 161-170.

C5. Helic, D., Strohmaier, M., Trattner, C., Muhr, M. and Lermann, K. 2011. *Pragmatic Evaluation of Folksonomies.* In Proceedings of the 20th international conference on World wide web (WWW 2011), ACM, New York, NY, USA, pp. 417-426.

C6. Trattner, C. 2011. *NAVTAG - A Network-Theoretic Framework to Assess and Improve the Navigability of Tagging Systems.* In Proceedings of the 11th International Conference on Web Engineering (ICWE 2011), Springer, pp. 415-418.

C7. Trattner, C., Körner, C. and Helic, D. 2011. *Enhancing the Navigability of Social Tagging Systems with Tag Taxonomies.* In Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies (I-Know 2011). ACM, 7-9 September 2011, Messe Congress Graz, Austria, pp. 18:1-18:8.

C8. Trattner, C. 2011. *Improving the Navigability of Tagging Systems with Hierarchically Constructed Resource Lists: A Comparative Study.* In Proceedings of the 33rd International Conference on Information Technology Interfaces (ITI 2011), IEEE, Cavtat / Dubrovnik, Croatia, pp. 173-178.

C9. Trattner, C. 2010. *QueryCloud: Automatically linking related documents via search query tag clouds.* In Proceedings of the IADIS International Conference WWW/Internet 2010, Timisoara, Romania, pp. 234-243.

C10. Helic, D., Trattner, C., Strohmaier, M. and Andrews, K. 2010. *On the Navigability of Social Tagging Systems.* In Proceedings of the Second IEEE International Conference on Social Computing (SocialCom 2010), Minneapolis, Minnesota, USA, pp. 161-168.

C11. Trattner, C., Helic, D. and Strohmaier, M. 2010. *Improving Navigability of Hierarchically-Structured Encyclopedias through Effective Tag Cloud Construction.* In Proceedings of the 10th International Conference on Knowledge Management and Knowledge Technologies (I-Know 2010), Graz, Austria, pp. 363-370.

C12. Trattner, C., Helic, D. and Maglajlic, S. 2010. *Enriching Tagging Systems with Google Query Tags.* In Proceedings of the 32nd International Conference on Information Technology Interfaces (ITI 2010), IEEE, Cavtat / Dubrovnik, Croatia, pp. 205-210.

C13. Trattner, C. and Helic, D. 2010. *Linking Related Documents: Combining Tag Clouds and Search Queries.* In the Proceedings of the 10th International Conference on Web Engineering (ICWE 2010), Springer, Vienna, Austria, pp. 486-489.

C14. Trattner, C., Helic, D., Korica-Pehserl, P. and Maurer, H. 2010. *Click, Click — and an Educational Presentation is Available on the Web.* In Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2010, Chesapeake, VA: AACE, 2010, pp. 3932-3938.

C15. Hasani, I., Portsch C., Trattner, C., Kacjan, B., Helic, D. and Maurer, H. 2010. *Using Wiki for Rapid Authoring of Huge Numbers of E-Assessments.* In Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2010, Chesapeake, VA: AACE, 2010, pp. 3046-3053.

C16. Trattner, C., Steurer, M., Kappe, F. 2010. *Socializing Virtual Worlds with Facebook - A prototypical Implementation of an Expansion Pack to Communicate Between Facebook and OpenSimulator based Virtual Worlds.* In Proceedings of the 14th International Academic MindTrek Conference: Envisioning Future Media Environments (MindTrek '10). ACM, New York, NY, USA, pp. 158-160.

C17. Maglajlic, S., Helic, D. and Trattner, C. 2010. *Social Networks and eLearning: New Model for Learning at Workplace.* In Proceedings of 32nd International Conference on Information Technology Interfaces, IEEE, Cavtat / Dubrovnik, Croatia, pp. 373-378.

C18. Trattner, C., Hasani, I., Helic, D. and Leitner, H. 2010. *The Austrian way of Wiki(pedia)! - Development of a Structured Wiki-based Encyclopedia within a Local Austrian Context.* In Proceedings of the 6th International Symposium on Wikis and Open Collaboration, ACM, Gdansk, Poland, pp. 1-10.

C19. Trattner, C. and Helic, D. 2009. *Extending The Basic Tagging Model: Context Aware Tagging.* In Proceedings of IADIS International Conference WWW/Internet 2009, IADIS, Italy, Rome, pp. 76-83.

C20. Helic, D., Us Saeed, A. and Trattner, C. 2009. *Creating Dynamic Wiki Pages with Section-Tagging.* In Proceedings of the 1st Workshop on New Forms of Xanalogical Storage and Function, held as part of the ACM Hypertext 2009, pp. 18-22.