COLLABORATIVE TAGGING:

FOLKSONOMY, METADATA, VISUALIZATION, E-LEARNING, THESIS

A Thesis Submitted to the College of

Graduate Studies and Research

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

In the Department of Computer Science

University of Saskatchewan

Saskatoon

By

Scott Bateman

Keywords: collaborative tagging, metadata, e-learning, information visualization, tag clouds, folksonomy, the World Wide Web, the Semantic Web

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Computer Science

University of Saskatchewan

Saskatoon, Saskatchewan, Canada

S7N 5A9

Abstract

Collaborative tagging is a simple and effective method for organizing and sharing web resources using human created metadata. It has arisen out of the need for an efficient method of personal organization, as the number of digital resources in everyday lives increases. While tagging has become a proven organization scheme through its popularity and widespread use on the Web, little is known about its implications and how it may effectively be applied in different situations. This is due to the fact that tagging has evolved through several iterations of use on social software websites, rather than through a scientific or an engineering design process. The research presented in this thesis, through investigations in the domain of e-learning, seeks to understand more about the scientific nature of collaborative tagging through a number of human subject studies. While broad in scope, touching on issues in human computer interaction, knowledge representation, Web system architecture, e-learning, metadata, and information visualization, this thesis focuses on how collaborative tagging can supplement the growing metadata requirements of e-learning. I conclude by looking at how the findings may be used in future research, through using information based in the emergent social networks of social software, to automatically adapt to the needs of individual users.

Acknowledgements

The creation of this document has reminded me that alone I would have accomplished extremely little were it not for the advice, direction, and teachings of supervisors, colleagues, friends, and family. It is to all of you I owe a great debt of gratitude for your time, patience, and caring over the last two and a half years.

I was extremely fortunate to have been able to have Dr. Gordon McCalla as a supervisor. From day one I was made to feel like an equal and given the freedom to explore my ideas, which has made the last two years extremely meaningful. I cannot imagine falling into a better situation than the one I had with you, Gord. Thank you.

I must also thank Chris Brooks who has been a mentor and a "second supervisor"; especially in the beginning when my questions seemed endless. Chris' amazing talent, interest and love of his work is infectious and I can't imagine life in the ARIES lab without him.

I have taken part in work with other outstanding professors who have all been selfless in sharing both time and knowledge, which has giving me incredible insight that I could not have gained anywhere else. Thank you to Dr. Carl Gutwin, Dr. Peter Brusilovsky, Dr. Jim Greer, Dr. Dragan Gasevic and Dr. Stephan Weibelzahl.

I would like to thank my committee members Dr. Carl Gutwin, Dr. Julita Vassileva, and Darlene Fichter. Thank you for reading my thesis and providing your comments which have given be both new prospective and a stronger thesis.

Thank you to all my fellow students and colleagues that I have worked with in Saskatoon, Dublin, Pittsburgh, Belgrade and Vancouver. I look forward to the future and hearing of all your accomplishments. In particular, I would like to thank the members of the ARIES and HCI Labs who have put up with my ramblings most frequently.

Finally, I would like to thank my family. Your support and sacrifice have made my ambitions possible and allowed me to find work that I love. My parents and parents-in-law have provided both financially and emotionally as we moved 4500 kilometers across the country to a strange new prairie place with their grand children. To my brother, you continue to provide inspiration with your daily achievements. To my wife, Christine, and my children, Mara and Mateo, you have blessed me with your support and love and allowed me to be shut away on countless evenings and weekends; not to mention an entire summer on the other side of the Atlantic. I can’t promise this won’t continue from time to time, but my goal is to be there when you need me and hopefully a little before.

Thank you all.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The World Wide Web, as it approaches its twentieth year, has unarguably grown to be the single largest source of information in history. The need for personal information organization has also increased, thanks in large part to the progression of the Web itself. However, personal organization in the new Web society has received only limited research attention when compared with other areas such as web-based information searching or commerce. While computer-based keyword organization technologies have existed for more than twenty years, it is only since 2004 that it has caught main stream attention through its employment on social software websites. Collaborative tagging provides a simple and effective solution for personal information organization, based on user created keywords [79]. However, since collaborative tagging has appeared through several iterations of deployment rather than through a scientific or an engineering design process, we know very little about how it may be applied in new situations and in new applications, and how it might be improved. Empirical studies of tagging systems reveal that despite their simplicity from a user's perspective, they are surprisingly complex and reveal broad social dynamics which beg to be more deeply studied [34][40][64]. This thesis explores three distinct directions looking at how collaborative tagging may be applied effectively in new applications and systems.

The rest of this chapter will define collaborative tagging, explore its history, provide intuition about why it is effective, and place it in terms of other methods of metadata representation. We also provide a description of seminal work in the field which describes the complex dynamics of this relatively simple method of self-organization.

Chapter 2 describes the interface for collaborative tagging systems, which may be a key, in part, to why collaborative tagging has risen to common usage so quickly. We present an in-depth study of the little understood idea of tag clouds which describes and summarizes collaborative tagging and has seen wide application on the Web as a general purpose information visualization technique.

Chapter 3 contextualizes collaborative tagging within the field of e-learning by outlining a series of potential research questions. We start the investigation of collaborative tagging in new domains, by outlining a pilot study of student created tags.

Chapter 4 presents a system created to make collaborative tagging applicable to e-learning. We describe the system in detail and outline the findings of an exploratory case study of the system's use by students.

Chapter 5 outlines the advances and work in a new area of research, called socio-semantics, which has formed around the notion of leveraging information from social software systems (such as collaborative tagging systems) for creating more structured forms of metadata. We describe how our work fits into the field and provide future directions for research in this new field.

Chapter 6 concludes by describing how the research experiences gained in the writing of thesis have shaped new intuitions, and describes how we will seek to apply them.

1.1 **About Collaborative Tagging**

Collaborative tagging, or simply tagging, is now common on social software websites for managing web-based resources. Such websites allow for the organization and sharing of digital resources, including: photos (e.g., flickr), bookmarks (e.g., del.icio.us), and blogs (e.g., Technorati). On these websites, simple keywords called tags are used to categorize the information on the site (such as the photos or bookmarks). The tags are usually unconstrained although some tagging sites do not allow spaces or other non-alphanumeric characters to be included in tags.

To materialize this description of tagging, the interface and tag authoring process of the most well-known collaborative tagging system del.icio.us is provided below. Figure 1.1 shows the author adding a website for a recipe to his bookmarks on del.icio.us. As he starts typing, the system suggests the most likely tags he could be adding. Below the vocabulary of tags created by the user is displayed, which allows the user to have a reference to maintain consistency in the vocabulary. Below, the system displays the most popular tags of other users who have already added the website to their bookmarks under the "popular tags" heading. By default the new bookmark and tags are shared with others using the system (which can be changed by clicking on the "do not share" checkbox). The notes field allows the user to add self-reminders or give more detailed comments to other users.

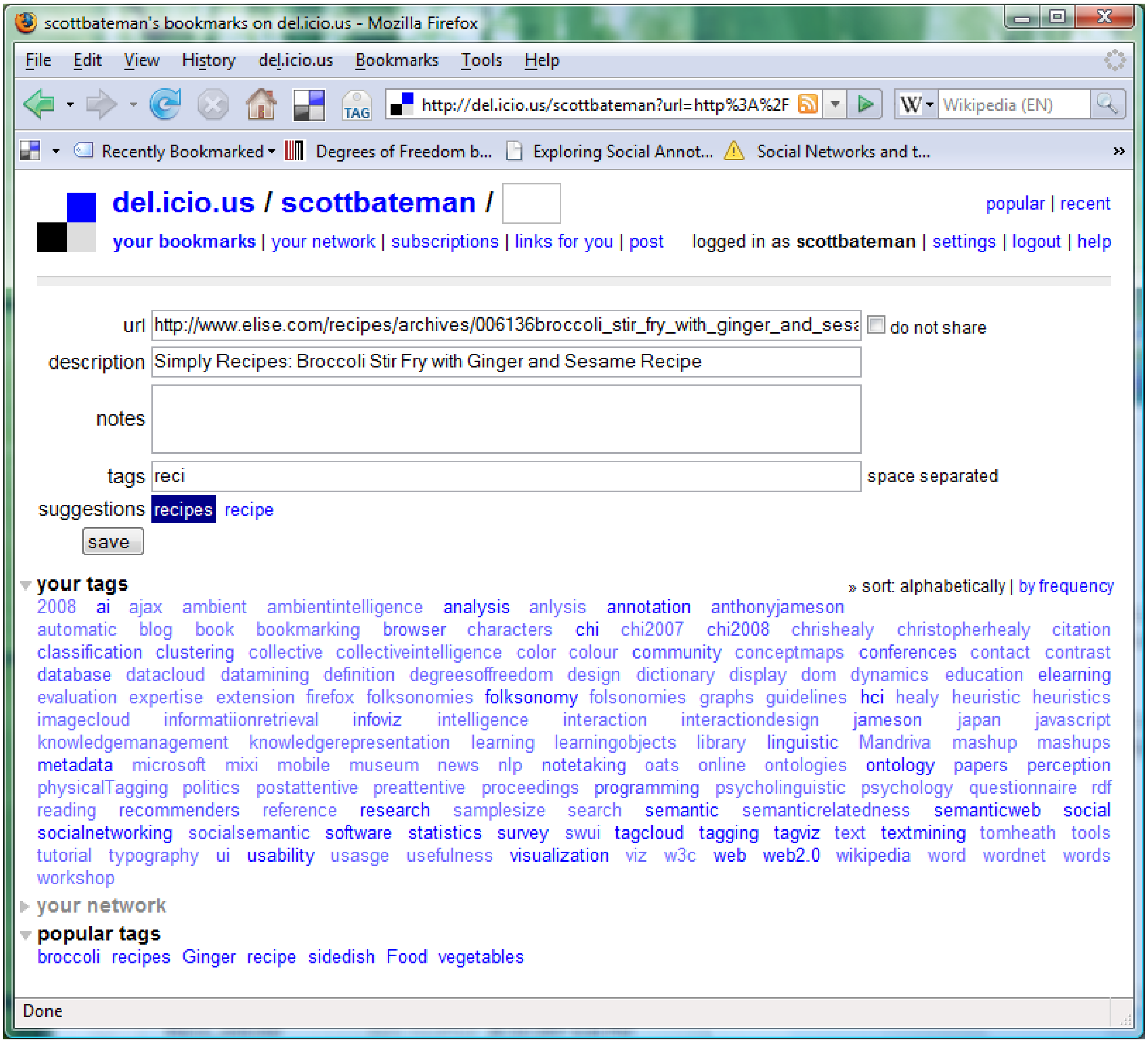


Figure 1.1 The author adding a website for a recipe to his bookmarks on del.icio.us.

Collaborative tagging is often incorrectly used as a synonym of folksonomy. Thomas Vander Wal, credited with coining the term, combined the word "folks" (multiple people with no particular designation) with "taxonomy" (a hierarchical structure of classification). Vander Wal defines folksonomy:

> … as the result of personal free tagging of information and objects (anything with a URL) for one's own retrieval. Folksonomy is created from the act of tagging by the person consuming the information… People are not so much categorizing, as providing a means to connect items (placing hooks) to provide their meaning in their own understanding… folksonomy is tagging that works…
>
> the three tenets of a folksonomy [are]: 1) tag; 2) object being tagged; and 3) identity, are [all] core to disambiguation of tag terms and provide a rich understanding of the object being tagged. [79]

So, while tagging is the simple act of adding keywords for personal retrieval while receiving community based support, folksonomy is more specific: it is the result of personal tagging; a vocabulary which characterizes a person-specific view of objects. Therefore, by viewing a folksonomy, including tags, tagged object, and user profiles, we are able to get a richer perspective than with simple keywords alone.

Collaborative tagging on the other hand is more general then folksonomy, although still relying on Vander Wal's three tenets. I extend the notion of the tenets of folksonomy to also include characteristics of collaborative tagging systems:

Sharing: the resources are shared (in most systems resources may be made private as an explicit option, and would then be only available to the owner and other owner designated users)

Collectivity: tags of all users accumulate on a resource and are available to all (on publicly shared resources)

Identity: user identities are incorporated to allow for disambiguation of tag meanings (through the exploration of the type of sites and other tags used by a single tagger, the context and meaning of all tags used by the tagger are more fully understood)

Support: recommendations of appropriate tags are made to users tagging a resource, and viewing particular tagger profiles can be useful (related to the concept of social navigation support[1])

The user interfaces of collaborative tagging systems are simple and lightweight (please see Chapter 2 for a more thorough discussion). Based on the characteristics of tagging systems (just described above) and the intuitive interfaces usually employed, bookmark-sharing tagging systems offer something which search engines do not: serendipitous discovery [58]. We define using a search engine as a finding or searching task, where either the user is looking for something associated with keywords (searching), or using keywords known to return a specific target (finding). Discovering resources on a social bookmarking site, such as del.icio.us, is a browsing process. Users click on words and users of interest in an attempt to find something of interest, in this case a target might be a new concept (represented in a tag), a new resource (such as a webpage), or a new user (e.g. somebody who has bookmarked a lot of interesting websites).

1.1.1 The History of Tagging

Lotus Magellan, a file organization program from the mid 1980's, allowed users to add keywords or tags to documents. Thomas Vander Wal reported watching a co-worker tag documents. The co-worker paid special attention to the files of others so that he could add his own contextual dimension. By 2000 Bitzi[2], organized URLs using a similar scheme; however, it lacked an important contextual dimension of user identity, and as a possible result the web site enjoyed only limited usership [79].

---

[1] See section 2.1.1 for description of social navigation support.
[2] See http://bitzi.com

Del.icio.us[3] was released to the public in 2003, and quickly became a very popular site for storing and sharing website bookmarks. This success was due in large part to the fact that users could access bookmarks between machines and browsers. However, it also incorporated a strong aspect of user identity, not available in previous systems, with an intuitive and lightweight interface.

Since 2003, tagging has been become an almost pervasive feature on social websites such as blogs and photos sharing sites. Tagging has also jumped from the Web to the desktop. Tagging of files is available in both Microsoft Vista (see Figure 1.2) and the latest version of Mac OS X, and has been incorporated in email clients such as Mozilla Thunderbird (see Figure 1.3).

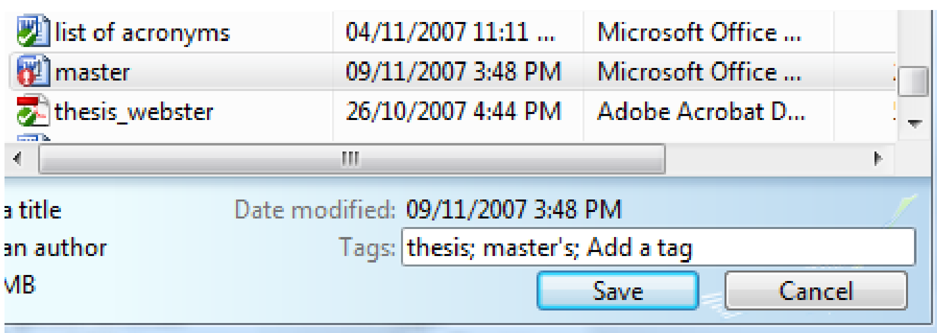


Figure 1.2 Tagging a file in Microsoft Windows Vista

---

[3] See http://del.icio.us

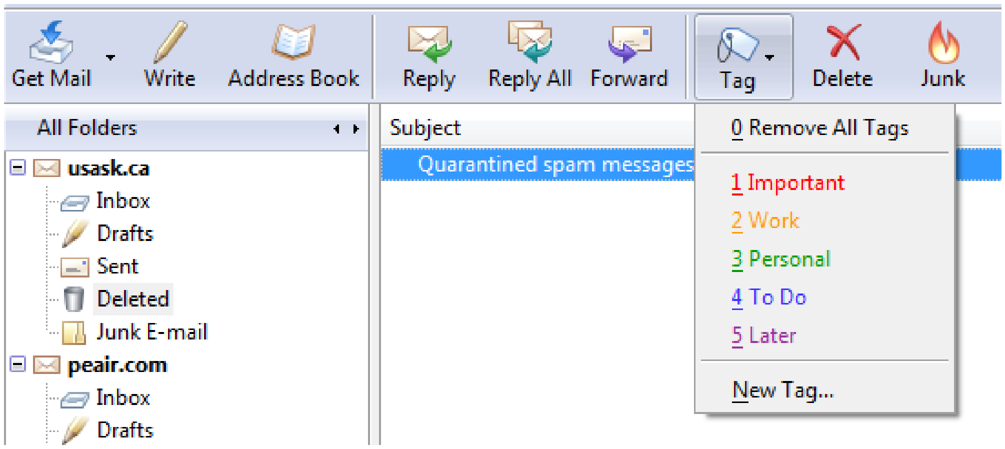


Figure 1.3 Tagging an email in the Mozilla Thunderbird email client.

1.1.2 Relationship to Taxonomies

Golder and Huberman describe tagging as being an open and inclusive organization scheme, as opposed to taxonomical or hierarchical schemes which are closed and exclusive. In well known taxonomical schemes such as the Dewey Decimal system, the objects, such as books, are placed within a single unambiguous category. Each category is a subcategory, where the parent category is more general and the child category more specific [34]. For example consider a taxonomy for the computer science domain. Java could be placed in the category of "object-oriented languages", while Lisp could be placed in the category of "functional languages". Both categories and languages could be placed into the higher, more general category of "programming languages". While the hierarchical structure in a taxonomy provides a lot of useful information about the domain it represents, it is limited in that it is usually exclusive. No one object can be classified in two spots on the hierarchy. For instance, Lisp could have just as easily been placed in the categories "interpreted languages" or "reflective languages", but many taxonomies impose an exclusive restriction where only one category can be chosen. Simple categories of items are rarely in real life mutually exclusive. Tagging in contrast is inclusive, because it makes no assumption about mutual exclusion, and it is open since it does not require a category to exist in a pre-defined vocabulary.

1.1.3 Relationship to the Semantic Web and Ontologies[4]

The Semantic Web is a vision whose most notable proponent is the inventor of the World Wide Web, Tim Berners-Lee. In his seminal 2001 Scientific American article, Berners-Lee describes in detail a hypothetical world where intelligent multi-agent systems organize and augment the lives of everyday people [9]. Making it all possible are machine readable descriptions called ontologies, which are based on first-order and frame logics. The notion of "semantic" is most often referred to as adding machine referenceable definitions to a three-part assertion: a resource, a property of the resource, and a value of the property. For example such an assertion could be made on this thesis: the thesis, has author, Scott Bateman. Each part of the assertion could be represented by a Uniform Resource Identifier (URI), which would point to a definition or representation of the assertion part. Often values involving a person will be defined as reference to the individual's website. Scott Bateman would be represented by the URI: http://www.cs.usask.ca/~ssb609.

Anecdotal experience from research experience in the field suggests that future of the Semantic Web as it was originally described is somewhat bleak. A huge amount of money and effort have been poured into Semantic Web technologies and research, and after roughly ten years, we seem no closer to reaching the original vision. Ontologies are very complicated, and

[4] See Chapter 5 for more a description the detailed relationships between collaborative tags and Semantic Web ontologies

this alone has resulted in low uptake of associated technologies by the general web community. This is compounded by the fact that complications arise from the complexity or impossibility of mapping together assertions from different ontological representations. This has proven to be a very difficult task, as even rigorously engineered ontologies of the same domain are rarely consistent.

This being said, the effort has not been wasted. Emerging from the vision have been new web technologies that are widely used - but mostly in a closed set of applications - to link together large and complicated data sets. Further, the Semantic Web community is undergoing a shift in focus from large, rigid and comprehensively engineered representations of domain knowledge for the web in general, to a somewhat more practical view. James Hendler, a leading Semantic Web researcher, cites the Web 2.0[5] movement as the major influence for conceding the need to move in smaller steps towards a Semantic Web of data, Hendler reminds researchers that, "a little bit of AI can have a big effect." [47]

1.2 **Tagging Dynamics and Emergent Characteristics**

In this section I briefly survey seminal work regarding the dynamic aspects and emergent characteristics in collaborative tagging systems. Recent work shows a tendency of tags from users to come together and stabilize in vocabulary (the tags used to describe a resource) [40]. Further, some research suggests that the popular tags are objective (fact-based) and may best represent the important terms in a community of interest [72]. I believe these desired tendencies of tags can be largely attributed to the interface and community support of tagging

[5] See section Collaborative Tagging Framed by Today's Web for a description of Web 2.0

systems. As the quality of notes in note-taking system has been shown to vary widely with interface capabilities [8], I believe the interface of tagging systems promote a healthy growth of tags.

1.2.1 Factual Tags

Sen et. al. performed a study on the quality of tags between several condition groups via the MovieLens movie recommender system [72]. For several weeks system users were provided with different tagging interfaces for organizing movies. The four condition groups of the study represented equivalent systems with the exception of the type of community-based recommendation of tags provided. The four conditions which presented recommendations for candidate tags to be applied to the movie were: no recommendations, random recommendations, collaborative filtering recommendations (using a cosine similarity algorithm), and popularity-based recommendations (providing tags that are the most popular for the given movie). The study found that the recommendations provided to users had a profound effect on the quality of the tags provided. The users in the popularity-based recommendation condition provided 82% factual tags, which are tags that represent an actual characteristic or property of the move (e.g. “Indiana Jones and the Temple of Doom” tagged with “Harrison Ford”). Other possible categories for tag classification used in the study were subjective (opinion-based) and personal categorization. This was much higher than the other conditions: collaborative filtering (67%), random (37%), and no recommendations (38%).

The results were important to show that the recommendations for tags were actually used by taggers. Intuitively one might assume that the collaborative filtering algorithm would perform the best, since it provides tags which are provided by others based on “similarity” between users, and would therefore represent some shared understanding. However, the most popular tags for a movie or any resource are tags that represent most strongly a consensus of all users and these tend to be factual.

1.2.2 Consensus: The Stabilization of Tags

Halpin, et. al, performed a large scale analysis based on data from del.icio.us. The study looked at the evolution of tag distributions over time on per resource (URL) basis [40]. They found that a power law distribution could be used to describe the formation of tags on a per resource basis. Further, they found that the tags representing sites on del.icio.us moved to such distributions relatively quickly[6] regardless of the number of annotators.

1.2.2.1 The value of the power law

Figure 1.4 shows the power law curve, which could represent an individual resource. In such a graph the y-axis would display the number of times a tag had been applied, while the x-axis would display the application of an individual tag sorted by frequency.

In this figure, the image is split into two parts, representing an 80-20 split. The "head" is represented in the left part of the curve, while the right represents the "long tail". This is based on applying the 80/20 rule of thumb made well known in Web circles by the working paper on Internet economics, "Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales" [16]. This paper describes how web stores like Amazon cater to the minority of a market, since the 80/20 rule dictates that 80% of business comes from

[6] "Quickly" was referred by Halpin et. al. as being over the span of a few months; they did not provide information on the mean tag applications or mean number of taggers before convergence.

20% of the customers. Amazon, with its giant library and without the constraints of a brick and mortar store is able to cater to purchasers in the long tail, with rarer books; accounting for 40% of all sales.

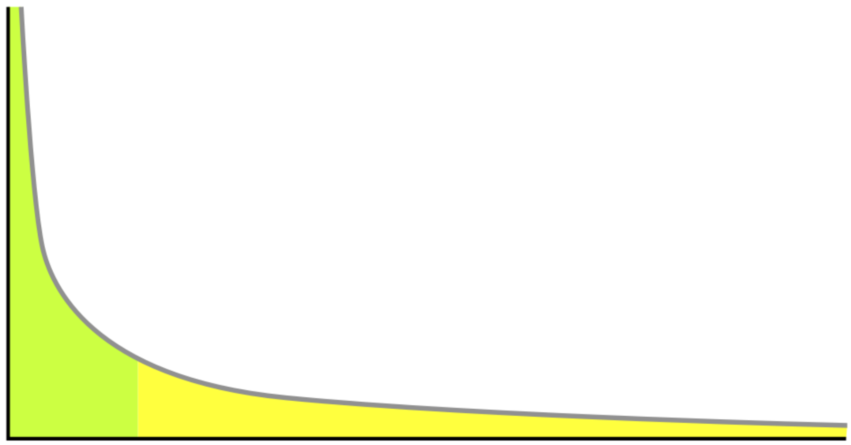


Figure 1.4[7] A perfect power law curve. Such a distribution can be used to describe the distribution of tags provided for a resource.

Shirky first applied the idea of the long tail to tagging to describe the value of diversity of the tags and taggers on del.icio.us [74]. The "long tail" of tagging while representing less popular tags caters to the many people who have a different take on the resource and its purposes. This being said, much of the tail may contain "meta-noise", which are: erroneous, misspelled, or out of context tags.

Beyond the conjecture of Shirky, that the rare tags in the tail have value, the identification of the formation a power law allows us to hypothesize two other characteristics based on the previous work. Firstly, since the power law distribution of tags has been shown to

[7] Picture by Hay Kranen, in Public Domain

emerge consistently over many different resources, it may provide a method to diagnose the "health" of a tag set representing a resource or the applicability of tagging to small community. Secondly, given a power law distribution of tags, we could identify with a relatively high degree of confidence tags which tend to be factual (as described in section 1.2.1 Factual Tags), as they should be in the head of the curve. While, these hypotheses are logical steps based on previous research, a controlled study would be needed to test them.

1.2.3 Social Networks in Collaborative Tagging Systems

Peter Mika's seminal paper titled, "Ontologies Are Us: A unified model of social networks and semantics" [64], has sparked much interest in using collaborative tagging as a more general solution for information organization. Mika was the first to analyze tagging data based on a tripartite relationship of actors (taggers), concepts (tags) and instances (resources). Using graphs representing the tag relationship, he produced networks by analyzing the co-occurrence of tags in del.icio.us. The result was a large tripartite graph of actors, concepts and instances. To analyze the relationships of the tripartite graph he reduced the graph into two bipartite sub-graphs: an actor-concept graph (where similar users and their concepts are represented) and a resource-concept graph (where similar documents and their concepts are represented).

Mika produced two concept sets for each of the two bipartite graphs, both of which represented the Semantic Web community. By interviewing the international Semantic Web community and through qualitative analysis he concluded the concepts in the actor-concept graph were significantly better to represent the community than the resource-concept graph. The social network model and its application have provided new perspectives on the emergence of concepts from a community at work. While the individuals in the community have only limited knowledge of each other, they are linked together by their actions rather than through artifacts.

Social networks are a way to discover perspectives of a community, which are represented in what has emerged as the most important concepts (in the tags themselves). This is because the actor-concept model accounts only for those connected individuals who share some conceptual common ground. Mika's results continue to be of significant interest to the Semantic Web community, as they look for more practical approached to best represent domains of knowledge. From the standpoint of collaborative tagging researchers the actor-concept model may provide a way to connect people through the tags they have used, provide more focused tag suggestions, or be a method to personalize search results.

1.3 **Collaborative Tagging Framed by Today's Web**

Collaborative tagging is among a new breed of simple yet effective technologies that have garnered both a lot of users and a lot of interest. Web 2.0 is a term trademarked by O'Reilly Media and used to characterize the "new version" of the World Wide Web which leverages the contributions of the casual web surfer to create, organize, and share knowledge. Among Web 2.0 technologies are Wikis, blogs and podcasts [67]. While there is much rhetoric and buzz around the new applications, technologies and websites, the ideology and paradigm shifts on the Web have been real. There has been a steady movement in web publishing from domination by business and media to an increased focus on individuals, their social groups, and their social interactions [37]. Mika has summarized the social connectedness on the Web today by saying: "We are not just building the Web any more: we are on it." [63]

Collaborative tagging among the new breed of Web technologies is unique in that it provides complex data that may be quantitatively analyzed. As such, tags represent little tidbits of knowledge that connect and represent people, and whose broader implications are continuing to be discovered.

Chapter 2

Navigating and Visualizing Tags

While the tripartite relationship of collaborative tags, individual taggers and tagged artifacts has revealed complex dynamics, the interfaces for navigating this information have remained relatively simple, not only to use but to implement, which has no doubt encouraged its wider uptake as an added feature for many websites. This chapter describes existing research about the tagging interface and presents a broad exploratory study seeking to understand the most popular method of visualizing tagging data, the tag cloud. In this study we sought to provide understanding as to how people make use of tag clouds and to provide directions for future research.

2.1 Navigating Tags… and Users

Pivot browsing is a term created to characterize the lightweight interactions that are common in tagging systems. Pivot browsing allows users to "reorient" their view of information, by clicking on different tag artifacts [62]. For instance, clicking on individual tags would give a list of resources annotated with the tag, or clicking on a username would provide a user's folksonomy (all the user's tags and tagged resources). A user can pivot on tags and users in combination to filter browsing results, for example clicking first on a tag, say "java", and then clicking on the username, "scottbateman", would allow the user to see all resources that scottbateman tagged with “java”. While these are not requirements they represent aspects of successful pivot browsing based interfaces. Pivot browsing has proven to be a lightweight and simple method to interact with tagging data, but it does not facilitate more complicated queries or provide answers to higher level questions of the data. For instance, finding the most popular tag describing a resource requires reviewing all tags and their number of applications sequentially, to find the tag with the highest score.

Pivot browsing itself is the feature which enables the serendipitous discovery of resources. This is a feature which is not available in search engines or in large directories such as Yahoo. Thus, the user interface of collaborative tagging has something very valuable to offer which is lacking in other methods of search and retrieval on the Web.

2.1.1 Social Navigation Support

Social navigation support (SNS) techniques are based on social navigation theory. Dourish and Chalmers define social navigation in information spaces as “moving towards clusters of people,” “selecting subjects because others have been examining them,” and “[driving] the actions of one or more advice providers.” [26] In short SNS techniques make use of past users’ interactions with a system to guide new users of the system, relying on the collective knowledge of the larger community.

Support can be provided directly or indirectly, and can be given either by a human or an artificial agent (in the role of an advice giver). Direct SNS would be the explicit communication of advice on navigation (e.g. a friend emailing a link and commenting, “you would be interested in this page”), while indirect SNS is discovered by examining the historical data traces left by other users [77].

The pivot browsing interface is well suited for navigating tags and maybe more importantly, navigating users. Millen and Feinberg have provided analysis of usage data in the Dogear enterprise tagging system. They discovered that more users had been browsing by clicking on user links than tag links. This suggests that the interface is equally well suited for navigating information beyond the tags alone. Further, it points to the use of indirect social navigation support, since user profiles advise on appropriate tags and resources [62].

While, pivot browsing itself has been a key to tagging popularity by facilitating data exploration, it does not summarize data for quick review, making more casual tasks such as impression formation and browsing a multi-step process. Instead at each step a user would need to navigate in and out of tags and user profiles. While these seem to be important tasks for tagging users, a more effective method of achieving these goals has been to use tag cloud visualizations.

2.2 **Tag Clouds: Visualizing Tags**

Tag clouds are visualizations of the content tags used in a website or other repository of information. The basic representation for the visualization is the tag words themselves; variables of interest, such as tag popularity or importance, are represented by manipulating visual properties of the words such as font size, colour, or weight (See Figure 2.1).

Figure 2.1 An example tag cloud from amazon.com/tags

Tag clouds have emerged along with the rise of collaborative tagging, and are now common on social software websites for photo sharing[8], bookmark sharing[9], and blog searching[10]. Tag clouds are frequently used as a way to give an overview of human-created metadata. Clouds have been shown to help people get a high-level understanding of the data, and to help people in casual exploration [71].

Unlike other information visualizations, tag clouds do not use artificial artifacts (such as the bars in a bar chart) to represent variables of interest. Instead, tag clouds merge the data variable with the data label (the tag word itself). Variables of interest are conveyed by manipulations of the visual properties of the text. The most common approach currently used is to map the popularity of a tag (its frequency of use) to the font size of the word in the cloud.

---

[8] See http://flickr.com
[9] See http://del.icio.us
[10] See http://technorati.com

However, many other possibilities exist both for input variables and for additional visual manipulations.

Tag clouds are a common feature on social websites, and have now been incorporated in many commercial websites as well[11]. However, very little is known about whether users actually interpret tag clouds in the way that the designers intend. There is a general understanding that words with very different visual characteristics will ‘pop out’ from the rest of the cloud [71][81], but less is known about how people interpret clouds in which there are no dramatic differences. In addition, it is not known whether there are interactions between different visual features: for example, in the tag cloud of Figure 2.1, ‘dvd’ is the word with the largest font size, but something about its visual characteristics make it less noticeable than other tags. In general, as tag clouds attempt to represent a wider range of variables with a wider range of visual manipulations, it becomes difficult to predict what will appear visually important to a viewer.

To find out more about how people interpret tag clouds, we carried out a broad exploratory study. Although there have been previous studies of information retrieval using tag clouds, our study explicitly looks at the interrelated visual properties of tags and how they convey a sense of importance to the viewer.

We created several sets of tag clouds that manipulated different visual properties (font size, tag area, number of characters, tag width, font weight, colour, intensity, and number of pixels). We asked participants to choose tags from these clouds that they felt were the most visually obvious, and used the results to determine which visual properties had consistent and strong effects. Our study showed several substantial differences in the way people used the

[11] See http://www.amazon.com/tags

different properties, and these results can used to improve the design of future tag clouds. However, we also found that when several visual properties are manipulated at once, the results become less clear, providing several starting points for future studies.

In the remainder of the chapter, we provide background on tag clouds, related work, and possible interpretation problems in tag clouds, and then introduce the visual properties that we studied. We then report on our exploratory study results, and discuss how our findings can be used by designers and researchers.

2.3 About Clouds

The first use of a tag cloud-like visualization was in Douglas Coupland's 1995 novel *Microserfs* [22]. Jim Flanagan's software module to provide tag cloud functionality for blogs was developed in 1997 and most likely inspired the photo sharing site flickr. It was flickr perhaps more than any other site or person which has popularized the wide spread use of clouds [17].

2.3.1 Clouds on the Web

Tag clouds, in the most common form found on the Web, are simply adjustments in font size of individual tags which reflect the popularity of the given tag (how often a tag has been used). However, there exist many other label-variable mappings, for instance: the recency of a tag's use encoded in color intensity on Amazon (the older tags fade away); or shared tag occurrences appear in red, in place of the normal blue tags, on del.icio.us. Swivel[12] presents data sets of interest, and uses the clouds as a general purpose visualization for any available data set.

[12] See: http://swivel.com

Such clouds are more generally called data clouds, since they don't display collaborative tagging data. The chainofthoughts website[13] provides an image cloud, where images grow as they receive more clicks.

2.3.2 What Tag Clouds are Used For

Tag clouds are simple and need little direction; in fact most tag clouds give little or no cues on how they should be used. For example flickr simply has the heading, "All time most popular tags", and presents a cloud[14]. This is followed by a description of what tags are, but no directions on how to use the cloud, or even that it is called a "tag cloud".

What is surprising given their simplicity is that tag clouds intrinsically support multiple uses. We list the tasks supported by tag clouds as suggested by Rivadeneira et. al. below [71]. We have altered the descriptions to be more generally applicable to tag clouds and data clouds:

Search: Locating or determining the absence of a specific target or an alternative target. Often as a means to get more detailed data related to the target.

Browsing: Casually exploring the cloud without a specific target or purpose, often drilling down on multiple discovered targets as they pique interest.

Impression Formation and Impression Presentation: The cloud can be scanned to get or give a general idea about a subject. Visually prominent items cause the greatest impression about the subject of a dataset, but other less prominent items also serve to enrich the impression.

Recognition or Matching: Recognizing the entire cloud as data which describes a subject.

---

[13] See http://chainofthoughts.com
[14] See http://flickr.com/photos/tags

2.4 **Related Work on Tag Clouds**

While tag clouds and data clouds are becoming increasingly prevalent most of their evaluation has appeared outside the scientific forum of peer-reviewed publications, but rather in the blogosphere [14]. Critics of note include information architects Jeffrey Zeldman and Thomas Vander Wal (known for coining the term folksonomy) [79]. Despite these subjective critiques of the tag cloud as a fad, Zeldman concedes, “[tag clouds] are smart, which is why so many have flocked to use them” [86].

Previous related work on improving tag clouds for information retrieval tasks have included semantically clustering tags [44], and altering tag placements and visual properties for more efficient spatial display [51]. While both works have provided interesting discussion neither has evaluated the effect on users. Below we present three relevant studies published on the nature and effectiveness of tag clouds.

Halvey and Keane studied the effectiveness of clouds for finding specific targets. They found that both vertical and horizontal lists outperformed tag clouds for finding a predetermined target. However they noted that targets with significantly larger font size were found more quickly. They also were led to believe that tag clouds were scanned rather than read sequentially, indicated by decreased target location times for tags in central positions [41].

The study of Rivadeneira et. al. consisted of two experiments. The first experiment examined the effects of font size, location, and proximity to the largest tag, on recall for sparsely populated clouds (clouds of 13 tags). They found the effect of font size was strong on recall, while proximity to the largest tag had no effect. A position effect was found, where the upper-left hand quadrant was strongest, but they attributed this to the sparseness of the clouds studied. The second experiment examined impression formation and recognition by varying the font size and layout of tags. Font size had a strong effect on recognition. They found no effect of layout on recognition, but found that it affected accuracy of impression formation. They found that a simple ordered list performed slightly, but significantly, better than tag cloud layouts [71].

Sinclair and Cardew-Hall examined the opinions and actions of participants who were given both a traditional search interface and a tag cloud for information retrieval. They found that for specific information retrieval tasks, users preferred the search interface. For more open-ended tasks, users preferred the tag cloud. They concluded that the tag cloud therefore is not sufficient as replacement for search in information retrieval tasks, but stands to facilitate the user's search process. The advantage of the tag cloud interface was summed up by one participant who said, "The tags [in the tag cloud] invoked ideas immediately in my mind, making access easier" [75].

2.4.1 Pre-attentive Processing

Pre-attentive processing theory proposes that changes can be made to symbols, such as altering colour or shape, which will make them "pop-out" to viewers. Pre-attentive processing theoretically occurs before viewers allocate conscious attention. Ware categorizes what can be pre-attentively processed into four major categories: form, color, motion and spatial position. All, with the exception of motion, are related to the visual features of tags in tag clouds [81].

The differences in features which do make things pre-attentively processed are usually quite stark, and it becomes more complicated as the number of visual features increase. Tag clouds always have more than one visual feature used at a time (at least two: i.e. position and font size). In addition, we hypothesize that people are very much "attentive" in selecting tags in a cloud. Although certain tags may pre-attentively pop-out, because of large visual differences, there should still be an attentive period of reading, reasoning and deliberation before a tag is selected.

2.4.2 Interpretation Problems

The problems with interpreting tag clouds arise from the interrelated visual features of text. To illustrate such problems Figure 2.1 presents the main tag cloud from Amazon. We informally asked 10 people in our research lab two questions about the Amazon cloud: 1) What is the most important word in the cloud? 2) And why did you pick that word? Of the 10 responses: 3 chose “music”, 3 chose “science fiction”, 3 chose ”fantasy” and 1 chose “history”. When asked why they had made the particular choices all made comments relating to size and colour intensity. Through examining the source code of the cloud, we found the largest tag to be ‘dvd’, which no one had selected. When told that ‘dvd’ was the largest tag a few said, “I didn’t even look at it.” Two others attributed their oversight to the fact that “it’s shorter.”

The Amazon cloud uses intensity to encode how recently a tag has been used; the older tags are faded. The World Wide Web Consortium (W3C) has provided contrast guidelines for information encoded by color. These guidelines allow people, including those with color deficiencies, to be able to adequately see all color-encoded information [21]. Further, it is known that the higher the level of detail required the greater the amount of contrast needed to convey the information, and sensitivity to contrast decreases with age [81]. We used a Colour Contrast Analyzer[15] which implements the contrast measurement algorithms suggested by the W3C.

[15] See http://juicystudio.com/services/colourcontrast.php

Based on results from the tool, we estimated that roughly 30% of the tags in the Amazon tag cloud (Figure 2.1) did not provide enough difference in brightness and did not provide sufficient difference between the background and font colour to be reliably seen by all viewers.

While, we know tag clouds do work in general, we cannot be sure that visual properties of tags are seen by everyone in the same way or in the way the designer intended.

2.5 **The Visual Composition of Clouds**

2.5.1 Visual Text and Font Properties

With the Cascading Style Sheet Level 3 (CSS3) working draft definition the W3C has defined a comprehensive array of properties in the text, font and colour modules [76]. Newer and more advanced properties, such as text embossing, remain largely unimplemented in today's browsers, but define what will be possible in the near future. When this huge list of visual text-related properties is available natively in web browsers many new avenues for textual encoding of information will exist.

We have investigated a smaller list of common visual text and font properties that are available in modern graphical web browsers. While the list is not exhaustive, it provides interesting properties that should be familiar to all people with web browser experience. We will use the term text to refer to both the properties of the text and font modules as defined by the W3C CSS3. Some visual text properties of note: text decoration (underline, over-line, line-through), text style (italics), text size (font size), text weight (bold), text colour, text intensity (contrast or saturation), text alignment (justification: full, left, right, center), character width variability (variable or mono-spaced), and word spacing (spacing between individual words).

2.5.2 Visual Characteristics of Words in Tag Clouds

In addition to the properties relating to font and text properties there are complex relationships between the visual properties of a tag in a cloud and the tags around it. These arise from the shape and visual features of characters as they make up a word, and how the words are arranged in a cloud. Such visual characteristics are: word length (the length in pixels), number of characters (in a word), number of pixels in a word (which relates to the hypothesis that: visually some characters stand out more than others, i.e. the letter 'm' is more visible than the letter 'i'), area of the word (as a word has different length), position in cloud, and effects of proximity (i.e. the visual effect of smaller tags being clustered around a dramatically larger tag) [71].

2.6 A Study of Visual Importance in Tag Clouds

To explore the effects of these various properties and characteristics of text, and to find out more about how people interpret tag clouds, we carried out a broad exploratory study. In the following sections we first present an overview of the goals and approach of the study, and then report specific methods, results, and interpretation.

2.6.1 Overview of the study

Our goal in this evaluation was to explore the effects of several different visual characteristics of tag text, in several combinations – with different numbers of interacting characteristics manipulated at a time. Therefore, the study was designed around testing multiple clouds with multiple manipulations. We identified ten configurations of the visual characteristics described below, and generated ten clouds for each of the ten configurations. In each cloud set any characteristics that were not under study were held constant (see Table 2.1).

Our main measure in the study was the degree to which users selected tags that had a given visual property. Since people might select some of these tags at random, we compared people's selection rate to the expected value (the rate that would be expected if people chose tags at random). We then could test whether the difference between people's selection rates was significantly different from the expected (random) rate. For example, if people chose 8 of 10 tags that were bold, and if only 20% of the tags in the cloud were bold, then we would have evidence that bold was used as a selection criteria.

This meant that for each tag cloud, we pre-identified a 'goal set' of tags that had each of the visual properties in which we were interested. We then used this set to score people's actual responses, and calculate our selection rates for each property. In cloud sets where there was more than one manipulation, it is possible that a single tag could have several properties of interest; in these cases, we tested against multiple goal sets (the interdependence between properties was recognized, and is taken into account in our interpretations below).

2.6.2 Study Methods

2.6.2.1 Independent variables: visual properties of text

Based on the visual properties of tags coupled with the font and text properties already described, we determined a list of interesting visual properties for study. For each property, the study investigated whether the property is chosen more often when people select 'visually important' tags.

1. Font Size

The most common manipulation in existing tag clouds is font size. Given the popularity of tag clouds it seems likely that there is strong correlation between font size and the interpretation of importance. However, we do not know how the other properties relate to size, and it is not clear whether size is always the most important cue.

2. Tag Area

Previous work by Kasser and Lemire has proposed using packing algorithms for more efficient display of tag clouds [51]. In order to manipulate tags in such a way to fit in the smallest area possible, they exchange the font size mapping with that of the area of the tag. This type of mapping only works if area is at least as effective as font size for conveying information.

3. Number of Characters

When we read, we recognize a word's component letters, and then use that visual information to help recognize words [55]. We were interested in the question of whether words with more or fewer letters attract our attention.

4. Tag Width

Words, even if they have the same number of letters, have different sizes. For example, the word 'illicit' takes up less horizontal space than the word 'snoozes' in a variable-spaced font, although both have 7 letters. Given that we recognize each individual letter when reading, perhaps wider letters and words are more visually important.

5. Font Weight

Changing the thickness of letter strokes, by using bold, is a common way people make words stand out in typography. We were interested in whether bold tags are more obvious than non-bold tags.

6. Color

In information visualization, color is a frequently-used visual feature for encoding data [81]. Tufte suggests that we can detect even a single pixel that differs in color on high resolution display [78]. How does a more even distribution of color influence what we consider important?

7. Intensity

The tag clouds at amazon.com couple font size with intensity; the W3C indicates that the higher the contrast of a word, the better for reading clarity [21]. We were interested in whether intensity affects interpretation of importance.

8. Number of Pixels

Letters have individual characteristics that may make them more visually prominent. Some letters take considerably more ink than others, and we were interested in the effect of the individual letters that make up a tag.

9. Position

If clouds are organized alphabetically, a tag's position does not indicate anything about its importance. However, people often start reading in a particular place in a document (e.g., the top left for English speakers). Are users therefore biased towards selecting tags that are in certain areas of the cloud? Do certain areas of a tag cloud get overlooked?

2.6.2.2 Interdependencies of visual tag features

The nine visual properties outlined above are difficult or sometimes impossible to study individually because of their interdependencies. Figure 2.2 shows the relationships of the visual features examined. The arrows show the other properties affected when a given property is changed.

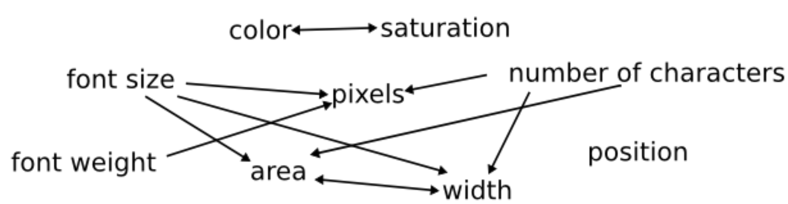


Figure 2.2: The interdependencies of selected visual features of tags.

2.6.2.3 Dependent variable

As discussed above, our study is focused on what visual characteristics led people to read and select certain tags in clouds. Our dependent variable in the study, therefore, is the degree to which people found each property to be visually important. The term 'important' was left as an open-ended concept in the study, to allow participants to make up their own minds about what visual properties are noteworthy. (Our use of importance should not be confused with algorithmically determined notions of 'interesting' tags [27]). As described above, what we actually measured in the study was people's selection rate of tags with particular properties, and compared these rates to the expected values (the rate that would occur if people chose tags at random).

2.6.2.4 Participants

We recruited 12 participants (8 men, 4 women) from a local university. Participants ranged in age from 18 to 30 years (mean 24). All used computers at least 28 hours per week (mean 46). All participants had normal or corrected to normal vision and could distinguish between all colors used in the experiment. Of the 12 participants 5 reported knowing what a tag cloud is, and 4 said they have used tag clouds before. We also asked of those people using tag clouds before to rate their familiarity, on a 10 point scale, with 10 being "Very familiar: I use them frequently" and 1 being "I have used them only once". The ratings familiarity ratings ranged from 1 to 8 (mean 3.25).

2.6.2.5 Apparatus

A custom study system was built in Java, which hid the underlying operating system (windows, taskbar, etc.) from view. The system presented to the participants a series of tag clouds, and recorded all unique tag selections (see Figure 2.3). Tag words for the clouds were obtained from the Kucera-Francis list [53]. The study was conducted on a Pentium 4 based computer, running Ubuntu and using an Apple one button mouse for input. The display was at 1280x1024 resolution on 19" 4:3 LCD flat screen.

Figure 2.3: A tag cloud in the testing system.

2.6.3 Tasks

The study presented a series of 100 tag clouds that were randomly generated before the study (139 to 147 tags per cloud). Each trial started with a blank screen that required the participant to click a button to load a tag cloud. The participant was given the instructions: "Click on 10 tags that you find the most important in the cloud. 'Most important' means the most important in relation to the other tags in the cloud. Please choose the tags based on how they look to you rather than the semantic meaning of the word." The system registered only the first 10 clicks and the participants were then required to finish the trial by clicking on a button outside of the cloud. There were 10 tag selections required for each of the 100 total trials (10 trials x 10 cloud sets). All participants saw the same clouds, though the order of presentation was varied by maintaining a consistent ordering but cycling through the starting cloud set.

2.6.4 Cloud sets

We tested 10 *cloud sets* – a group of tag clouds containing the same independent variables. Each of our cloud sets contained 10 clouds, and manipulated a different set of visual properties (see Table 2.1); due to the interdependencies between properties, some properties could not be tested alone. All other visual features were fixed for each group. Tag positions were random over all trials, and were explored only in secondary analysis. Table 2.1 shows the independent variables by each cloud set.

The cloud sets can be organized into three main groups: those with one manipulation (cloud sets: 1 and 2), with two manipulations (cloud sets 3-6), and real-world-style clouds with several manipulations at once (cloud sets 7-10). Cloud set 7 was a basic multiple-manipulation cloud with equivalent tags; cloud set 8 was based on flickr and del.icio.us; cloud set 9 was based on Amazon, and cloud set 10 manipulated all variables excluding pixels.

	Tag Cloud Set									
Visual Property	1	2	3	4	5	6	7	8	9	10
Font Size							X	X	X	X
Tag Area							X	X	X	X
Num. of Characters			X					X	X	X
Tag Width	X		X				X	X	X	X
Font Weight				X	X					X
Color						X				X
Intensity					X	X			X	X
Num. of Pixels		X		X			X	X	X	

Table 2.1 The independent variables for each experiment.

When visual text features were not being examined in a given experiment, their values were not visibly different. Table 2.2 shows the range of values used for each visual feature when used as an independent variable. The ranges of variables selected were somewhat smaller than what is typically found in tag clouds; this was done in order to avoid more 'extreme' pop-out tags.

Visual Property	Value Range
Font Size	26-36 pt.
Tag Area	Dependent on font size and number of characters.
Num. of Characters	3-7 characters
Tag Width	Dependent on font size, spacing and # characters.
Font Weight	bold or normal
Color	blue or red
Intensity	100%, 87.5%, 75%,62.5%,50%
Num. of Pixels	tags varied over a 300 pixel range

Table 2.2 The value range for independent variables.

2.6.5 Procedure

Participants were introduced to the study, given a demographics questionnaire, and randomly assigned to one of the ten ordering groups (which varied the presentation of cloud sets). Before starting, the participants were also shown three example tag clouds from prominent websites, but were not given any instruction on how they should be interpreted. After the first 3 trials in each condition users were asked to state their selection strategy – that is, to explain why they had made those particular tag selections, via a popup text dialog.

2.7 **Tag Cloud Study Results**

We were interested in two aspects of the data: reliability and strength. Reliability was tested using statistical analysis which compared the selection rates to the expected rate (since our data was not normally distributed, we converted rates to ranks, and carried out a non-parametric test). Strength was examined as a post-hoc investigation of the original means – for those properties that showed reliable differences, we inspected the means to interpret the size of the difference between actual selection rate and expected rate.

The study gathered 10 tag selections for each cloud. There were in 10 clouds in each of the 10 cloud sets (10 selections x 10 clouds x 10 cloud sets = 1000 selection per participant). Selection rates for each property of interest for a particular cloud set were converted to rank data, and compared against the expected value using a Wilcoxon Signed Rank test. A summary of reliability and strength results are shown in Table 2.3. Each row represents a visual property of interest, and the columns represent the reliability and strength results for that property. Each property was considered over all tag sets where it was a variable of interest (see Table 2.1 for this mapping).

	Strength of Effect		**Reliability (Wilcoxon)**	
Tag Visual Properties	**Mean Selected**	**Mean Expected**	**Z**	**p**
Font Size	8.58	2.03	-5.88	< .001
Tag Area	1.97	0.76	-5.11	<.001
Num. of Characters	2.65	3.63	-5.12	<.001
Tag Width	2.45	2.35	-1.97	<.05
Font Weight	8.19	3.26	-5.51	<.001
Colour (blue before red)	5.16	4.93	-0.37	>.05
Colour (red before blue)	4.84	5.07	-0.37	>.05
Intensity	5.96	4.00	-6.07	<.001
Num. of Pixels	1.33	0.70	-6.29	<.001

Table 2.3 Results from the experiment. These are summarized results where, all relevant cloud sets have been considered together for each visual property.

2.7.1 Participant Strategy Results

Responses to the free-form text field which appeared after each of the first 30 tag clouds are shown in Table 2.4. Each cell represents the count of cited visual properties in the participant responses, for each clouds set. If the participant stated that they chose at random or some other strategy which could not be reasonably mapped to an applicable visual property we added to the count in the 'Other' row. Only the cited visual properties are shown (number of pixels and tag area were not cited), and visual properties which were not varied in a cloud set have been grayed out.

	Tag Cloud Sets									
Tag Visual Properties	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Font Size							25	24	24	10
Font Weight					24	21				15
Color							23			16
Saturation						8	14		10	6
Position	7	1	5	4	2	2	2	6	5	2
Num. of Characters			3					0	0	0
Tag Width	1		3					0	0	1
Other	25	28	25	6	7	5	7	7	1	1

Table 2.4 The responses of participants as to why they chose certain tags.

2.7.2 Interpretation

1. Font Size

Larger font sizes were chosen reliably at the highest rate among any visual features. This suggests that even over small changes, of say a font size point or two, users can see differences in tags with a relatively high accuracy.

In cloud set 10 the rate of selecting large tags had gone down by roughly 3.5 selections (mean 5.917) compared to the other groups. This affects both the strength of the effect and reliability scores, and it can be attributed to a couple of factors. First, reviewing the responses of the participants collected on reasoning for selections revealed that only 10 times was size used as a basis for selections while font weight (bold) was cited 15 times and colour 16 times. This fact is coupled with the composition of the cloud set 10 clouds, in which the number of pixels in each tag was fixed. This resulted in bold words tending to be in smaller font sizes, as the words with a larger font size tended not to be bold. This is somewhat confounding for the results of group 10, since the distribution of visual features was not even. However, it also suggests that users were more apt to select tags with higher font weight values and based on colour rather than font size.

Overall, size seemed to be the most important factor when used, with the exception of cloud set 10, in which font weight and colour were most likely the important factors.

2. Font Weight

Higher font weight (bold tags) was selected reliably and at high rate above expected. When considering cloud set 4 which had only font weight and the number of pixels in tags as independent variables, based on participant responses, only 5 times did users not select based on font weight. One participant did not seem to notice or attribute importance to bold at all, over all 3 recorded responses, and selected just better than randomly choosing tags. All other participants selected bold tags at a very high rate, but not perfectly. This is despite having responded that they selected based on bold tags. It is impossible to know if they were distracted from their goal by some other alluring visual feature or selected incorrectly.

Font weight was the second highest selected property, and the responses seemed to follow these lines. However, in cloud set 10 where both font size and weight were used. Font size had a mean selection of only 1.433, while font weight was selected on average 6.241; this is higher than any other visual property. These numbers are supported in the responses where font weight was cited 15 times compared to 10 times for font size. This indicates that participants tended to choose bold tags before choosing tags with large font sizes.

3. Colour

In calculating results for colour we used two orderings for the top coloured tags. As half the tags were blue and half were red, we did not know which users would choose. Neither did our population, as neither ordering was selected reliably. This, however, was judged over all participants. We can see in both our data and our responses from participants, that there is almost a 50/50 split, those who chose blue and those who chose red. If we were to divide the group in two, those who chose blue and those who chose red, we undoubtedly would see a higher level of reliability. Colour when used was cited on par with or more important than font size. The reasons people made colour-based choices differed greatly: some chose red because "Red stood out.", while another said they ignored the colour because, "red burns the eye."

The use of both colour and intensity together seems to create interesting issues. We thought participants would always choose the darker (or high intensity) colours. However, some participants said they went for the "pinker" tags over either red or blue (tags with high intensity). Pink tags were actually red tags with lower intensity.

While we continue to believe that colour is an important visual property, our open ended task, of “selecting the most important tags”, did not lead to consensus in our population over what colour (blue or red) was more important.

4. Intensity

Higher values of intensity were selected reliably, but only slightly above the random. When intensity was cited, participants most often chose darker words, but a few also indicated the choice of lighter words or pinker (red low-intensity) words. It seemed people had no problem in noticing the intensity differences. We provided 5 levels of intensity of colors between 50% and 100%. Although people often said they chose darker tags, we don’t know if they set some arbitrary level of darkness (intensity) for their selections. For instance they could have chosen 80% intensity and higher as being ‘dark enough’.

We believe that higher values of intensity are fairly reliable in drawing the attention of participants. Although, we have shown evidence that in some situations lower values of intensity can also be interpreted as more important than higher values. This may be in-part to the overloaded nature of low-intensity red appearing to be a third color ‘pink’.

5. Number of Pixels

Tags rated with the highest number of pixels were selected reliably but not much more often than the random case. Variance in pixels was high, within 25% on either side of the average case. While the numbers indicate an element of reliability, it is more likely due to a correlation with other visual features, such as font weight and font size. Had there been a visual influence by increased number of pixels in words, we would have seen it in cloud set 2. This experiment only used the number of pixels in a word as an independent variable. The different features of the font and characters dictated the variance in pixels. While, some participants still said some words seemed to “stand out”, this was not correlated to the number of pixels ($Z = -1.081$, $p > .05$). So, it is fairly safe to assume that any effect based on the influence of pixels is due to its correlation to more prominent related visual features. Thus, very little visual importance seems to be attributed to the number of pixels in a tag.

6. Tag Width

Wider tags were selected only slightly better than random, and with relatively low reliability. As such this characteristic does not seem to have much weight in the interpretation of importance. Cloud set 1, used exclusively a variable spaced font as a visual feature; resulting in different tag widths. However, this allowed for only very small differences in tag widths. The maximum difference in any of the tags’ width was only 27 pixels. In fact tag width was only cited once as the basis for selections. This indicates that small differences in width have no visual importance.

7. Number of Characters

Tags with a larger number of characters, were reliably selected less than the random case. In cloud set 3 only the number of characters and therefore the width of the tag were independent variables. However, only 3 recorded answers noted that they selected tags of higher character counts. Also, in cloud set 10, bold tags where selected at the highest rate. In this cloud set bold tags tended to be shorter words (see interpretation of Font Size), which would reduce the mean of all tag clouds sets. Therefore, we do not attribute selections results to a preference for smaller tags. Overall people took no note of subtle differences of character counts, of say a single character, in their choices of tags. However, anecdotally in our informal experiences we have seen several times with large differences of characters, 3 or more, that the number of characters has been cited as a deciding factor in selections.

8. Tag Area

Tags with larger areas were selected reliably higher than the random case. We attribute this not to the strength of the effect it has but because it is correlated to other more prominent visual tag features. For instance, changing either font size or the number of characters would change the area of the tag. While some participants simply chose tags that were "bigger", we took this to be shorthand for bigger font sizes. Since, subtle differences in tag width went unnoticed we would hypothesize that this would scale to tag area as well. We do know that people consciously attribute importance to font size, and thus it seems unlikely that they would do the same for tag area. This was supported by the responses in which participants did not cite tag area once as a basis for selection.

2.7.3 Analysis of Position

We present position as a secondary result, as it was not directly studied in our design. We used clickmap visualizations to display all the tag selections by participants to get a qualitative view of position and its relationship to the tags in the different cloud sets. The clickmaps were normalized over all individuals for each tag cloud group. Figure 2.4 shows two of the click maps from two of the tag cloud sets, where lighter areas received more clicks than darker areas.

The tag cloud labeled cloud set 2 represents the 10 clouds where only the number of pixels was varied. This resulted in a tag cloud where all the tags were very similar. The dominant visual feature of the clickmap is the central tendency of the clicks. The cloud set 10 clickmap instead had 7 visual features as independent variables, including the stronger visual features (font size, font weight and intensity). The click maps presented exemplify two dominant trends that are easily noticed over all the tag cloud clickmaps. First, it seems the central tendency of clicks is reduced as the strength and/or number of visual features increases. This may be due to the reduced visual similarity between tags, requiring users to venture out from the center to find visually 'important' tags. Second, the areas represented by the top and bottom rows remain under used by the participants.

In the response collected (see Table 2.4), participants cited position in every cloud set, but not at a high rate. We can see no correlation from the tag cloud sets to the number of times position was cited. It would seem participants did feel position had a limited role in visual importance.

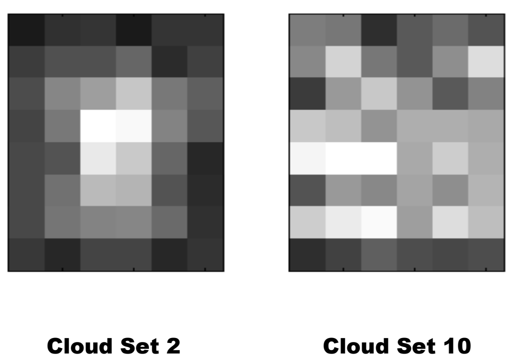


Figure 2.4 Clickmaps displaying the concentration of clicks for the clouds in two of the ten tag cloud sets. Lighter areas received more clicks than darker areas.

2.8 **Discussion of Study**

2.8.1 Summary of Results

Overall, our results indicate visual properties that are easily seen (such as font size, font weight, saturation and color), are likely candidates for interpretation of importance. Features that are not easily visually assessed (such as number of pixels, width and area) are less likely to be interpreted as important. Colour seemed to be easily distinguished, but it was not clear which colours are more visually important. Centrality of selections seems to be influenced strongly by the number of visual properties used. Even though position had a limited role overall, there seem to be areas that have less influence (top and bottom), suggesting the center is a safer zone for prominence. Finally, we have also seen that small differences in characters (1 character) have little influence on interpretation of importance, although we have seen anecdotal evidence of large differences in characters serving as a visual tie-breaker.

2.8.2 Lessons for Designers of Tag Clouds

Below we summarize our results for each visual property we explored, and state advice that we believe designers of tag clouds can readily apply.

2.8.2.1 Important visual properties

Font Size was consistently recognized in our study, and people were able to recognize even small changes in size; it is therefore appropriate for data with a large range of values. However, when it is used with other more important properties, it may lose some of its visual significance.

Font Weight was consistently identified as an important visual factor. We suggest that font weight may be appropriate for binary-type data or to highlight tags that should be 'important' regardless of their other visual properties, as it seems to outweigh other visual features when used selectively.

Intensity provided a relatively good method to capture visual importance. We do not have firm indication of what are the smallest intervals that can easily be interpreted, but we do know that levels of roughly 10% difference seem appropriate. Designers should also be wary of lowering intensity too much, as it may provide problems for some viewers.

2.8.2.2 Less important visual properties

Number of Pixels, Tag Width, and Tag Area all have a strong correlation to other important factors; however, they were not identified in actions by the participants as being a visual feature of note. We feel that tag cloud designers can largely ignore these properties when designing clouds.

2.8.2.3 Visual properties to use with care

Colour is identified easily by viewers (with normal colour vision), although our study did not reveal to what degree and which colours are more likely to draw attention of a viewer. For this reason, if colour is to be used, an appropriate colour palette should be displayed with clear mappings for the viewer, and colours should be distributed evenly to avoid undesired pop-out effects.

Position was not measured directly by our study. We did see a strong inverse correlation between the number of properties manipulated and the centrality of tag selections. Designers should avoid a catch-22 in trying to avoid effects of centrality by incorporating more visual effects, since we have seen evidence that doing so reduces the effectiveness of all visual properties. More study is needed to see what role position plays in tag clouds; however, designers may want to keep tags they wish to be more prominent closer to the center of the cloud, and avoid the top and bottom lines. This may be difficult when typical alphabetic ordering is used.

2.8.2.4 Applicability of prototyping

Tag cloud designers may be well served to do small-scale prototyping when using new data-to-visual-feature mappings. These can be fashioned after our informal lab study using paper prototypes, and ask simple questions such as, "Which tags are the visually most important and why?" The answers will serve to inform designers of the appropriateness of mappings, and if the appropriate tags are accurately capturing the visual attention of their audience.

2.8.3 Tag Cloud Study Paradigms

Our study followed the simple paradigm suggested by Rivadeneira et. al. for tag cloud evaluation [71] and offers a variation which may be more appropriate for visual property analysis. In the original paradigm, there is first a cloud presentation period, the length of which is determined by the predicted usage, based on the task(s) or aspects being evaluated. The cloud presentation is followed by an elicitation period, where reflections are gathered. Our study provided a time-limit-free presentation time, which encouraged participants to scan clouds in order to finish the tasks in a timely manner. The main difference between our design and that of the Rivandeira et. al. study was that our primary data collection took place during the tag presentation, based on actual tag selections performed by users, instead of during the elicitation period (although eliciting the opinion of the participants also proved to be important). We also feel that ranking visual features may be an appropriate alternative to form-fill responses. We believe our design is more suitable for the questions we pose on the visual importance of tags, and will scale more appropriately for future endeavours using eye-tracking [41] [71].

2.9 Summary

Our exploratory study is the first to identify what visual features of tags draw the attention of tag cloud viewers. We have been able to make some immediate recommendations to the designers of tag clouds while providing a basis for future research with regards to understanding how clouds can be more effectively displayed. We have also described an extended study paradigm to better suit both questions about visual tag characteristics and logical next steps for evaluation using eye-tracking.

Future work could explore a more limited feature space based on our results, eliminating those features which are of limited importance. Using an eye tracking system would allow more immediate questions to be answered, in particular with regards to properties of colour and position. Further, we wish to leverage our results to explore how varying visual properties of tags affects various tasks using tag clouds.

In terms of the larger picture as a tagging interface, we have shown that tag clouds do effectively convey information to their viewers based on the most commonly used mappings. Along with pivot browsing, tag clouds also facilitate indirect social navigation support, by providing advice on what tags could be the most important. While some debate still remains about whether the tag cloud will persist as a common visualization technique, we feel that in terms of tagging it is an appropriate complement to pivot browsing interfaces. As such it is likely that tag clouds will persist as long as do current tagging systems, if not as a more general purpose information visualization technique.

Chapter 3

Applying Collaborative Tagging (to E-Learning)

In this chapter we investigate intuitions on the suitability of collaborative tagging to a new context: e-learning. We describe why tagging seems an ideal fit for e-learning research and outline a small-scale study based on collection of data from students in a university course using online learning content. The study gives some initial insight on the applicability of collaborative tagging for e-learning from the perspective of both the instructor and the learner. We then discuss alternative tag collection methods to overcome the cold start issues that would exist in new and small social software communities, similar to those in an online class. In doing so, our study and analysis illustrates a qualitative approach for assessing the applicability of tagging in a new domain and provides questions of interest for future research in applying tagging to e-learning.

3.1 Applicability of Collaborative Tagging to E-Learning

Our interest in investigating tagging in e-learning revolves around five main points. We believe tagging provides possible solutions to, or has interesting implications for, each of the points outlined below.

1. Organization and Annotation

Learning management systems (LMSs) currently lack sufficient support for self organization and annotation of learning content. Learning Management Systems such as WebCT[16], Blackboard[17], Moodle[18], and Sakai[19] provide only rudimentary methods for self-organization and content annotation, if any. However, walk through a university campus and you will see students engaged in a number of different organization and annotation activities. These include writing notes, creating marginalia in books, highlighting text, creating dog ears on pages or bookmarking pages. How can we best provide these traditional capabilities to students using digital learning materials? Are there methodologies better suited for such tasks in an online environment?

During lectures as many as 99% of students take notes [68], and 94% of students at the post-secondary level believe that note-taking is an important educational activity [83]. In this sense we liken tagging to note-taking, since tags represent an aspect or cue to be used in the tagger's recall process. Tagging provides a lightweight scheme for self-organization and most tagging interfaces provide fields for longer sentence based notes to be taken (as previously seen in Figure 1.1).

2. Metadata Collection

The learning object paradigm suggests that online educational content can be collected, aggregated, and packaged for delivery to learners. Key goals of this paradigm are to enable the

[16] See http://www.webct.com
[17] See http://www.blackboard.com
[18] See http://moodle.org
[19] See http://sakaiproject.org

discoverability, modularity, and interoperability of learning resources [32] [83]. However, the main hurdle in achieving this vision is the lack of meaningful and usable metadata describing learning objects. Collaborative tags represent a form of practical metadata, which could supplement in part the needs for detailed learning object descriptions (Please see Chapter 5 for a detailed discussion of learning objects and issues in metadata).

3. Implications of Social Software for Learning Theory

We exemplify learning theories by describing the implications of social software for a particular theory, constructivist theory (a more complete description of learning theory is outside the scope of this writing). Constructivist learning theory is a very widely researched area of educational theory. Constructivism, very broadly, proposes learning as a reconstruction of knowledge rather than as a transmission of knowledge. This reconstruction is usually stimulated in a learner by an expert (the teacher), who observes and challenges students' beliefs in what they know [69]. Control usually remains with the instructor. Employing social software may require relinquishing instructor control. For example learners can watch a knowledge-base of social navigation support[20] emerge which offers them advice on what is important in a course (e.g. learners could observe tag clouds describing course concepts). More research is needed to understand how tagging and other social software technologies may impact these and other learning theories in practice.

[20] Social navigation support (SNS) is described in section 2.1.1.

4. Knowledge Gain

Tagging, by its very nature is a reflective practice, which can give students an opportunity to summarize new ideas while receiving peer support (through indirect SNS). Tagging represents an action of reflection, where the tagger sums up a series of thoughts into one or more summary tags, each of which stands on its own to describe some aspect of the resource based on the tagger's experiences and beliefs [50]. Intuitively, when analyzed in terms of the classical Bloom Taxonomy of Learning [10], learners who use tags could show evidence of moving up the hierarchy from the lower "consumption"-based levels of learning (knowledge and comprehension) to higher levels of applied and meta-cognitive knowledge (application and analysis). Further, reviewing of tags (i.e. comparing tags used by a community of taggers) would potentially facilitate the move to the highest levels of Bloom's Taxonomy of Learning (synthesis and evaluation).

While the use of tags to encourage knowledge gain is hypothetical, the difference in knowledge gains afforded by alternative note-taking tools has been demonstrated by Bauer and Koedinger. Their study used a number of different interfaces for note-taking in online learning and they showed that there were significant impacts on knowledge gain based on the note-taking interface used [8]. Further study is needed to understand the implications and benefits of providing students with tags and how they would compare to other interfaces for organization and for impacting knowledge gain.

5. Pedagogical Reflection

In e-learning and distance learning there is a lack of the social cues that inform instructors about the understanding of new concepts by their learners. Collaborative tags, created by learners to categorize learning materials, would allow instructors to reflect at different levels on their students' progress. Tags could be examined at the individual level to examine the understanding of a learner (e.g. tags that are out of context could represent a misconception), while tags examined at the group level could identify the overall progress of the class. Working with instructors of online courses employing tagging would help shed light on the perceived benefits of reflection based on tags.

3.2 **Pilot Study of Tagging for E-Learners**[21]

In this section we explore the results of a pilot tagging study done on delivered educational content. While, the previous section outlined possible research questions based on the application of collaborative tagging in e-learning, we wished to first assess collaborative tagging in a new and smaller online community. The goal of the pilot study was to first assess how learners would apply tags and to learn of any possible e-learning specific issues.

3.2.1 Pilot Study Description

We asked two senior graduate students with teaching backgrounds to create short pieces of content relating to topics on artificial intelligence to be delivered to first year non-majors. These graduate students could write in any style they deemed suitable, but were given rough guidelines as to the form content should take (e.g. page length, reading level of students, and specific subtopics that should be taught). The content was then smoothed over for consistency by a single editor. We then collected tag and tag phrase sets using three different approaches. The

---

[21] This section contains content from [6].

first approach was by having the graduate students create (independently) a set of tags for each page of content (five pages in total). The second approach was to use the Automatic Metadata Extractor application [42] that has been designed specifically for creating learning object metadata in the form of Dublin Core or IEEE Learning Object Metadata standards. In our analysis, we have focused on the keywords being extracted, since sentences pulled out of the documents by the tool seemed relevant but not precise enough for use in tag clouds or keyword-based searching. Finally, in the third approach we asked students (over 200 unique participants, although not all students completed tags for each page) who viewed the content to "...write several keywords to describe this page (e.g. keywords you would use in a Google search to find this page)", and collected their input tags as well. Students were extrinsically rewarded for participating in this study by means of an entry in a draw. Only a short description of tagging was provided to students to allow them to apply keywords as they saw fit and to avoid overloading them.

3.2.2 Qualitative Results

A reading through the different tag sets created provides some interesting insights. The tags created by the Automatic Metadata Extractor were simple and included no tag phrases. These tags leaned towards very broad concepts. For instance, the tags created for the content around “Case Based Reasoning” included “case”, “reasoning”, “cbr”, and “problem”; all very high level. The instructor tags tended to provide very specific sub-topics that the page discussed. In addition, there was mixed use of both single tags and tag phrases, the latter of which provided more contextualized tags (e.g. “computational intelligence”, “human intelligence”, or “artificial intelligence”). Student created tag sets had a less than perfect signal to noise ratio, as some students wrote seemingly useless tags (less than 5% of all student tags were in this category). We believe this is due to low motivation – because of a low likelihood of being drawn for the extrinsic reward. However, most tagging systems are collaborative in nature, and thus intrinsic motivation can play a strong part (i.e. the learners in our study didn't necessarily view themselves a benefiting from the tags that were created).

3.2.3 Quantitative Results

After normalizing the tags[22], we plotted the number of tag occurrences against the set of all tags collected (using only the data we collected from the students) which is shown in Figure 3.1.

The students produced a total of 211 tags, of which 122 were unique. The experts produced 60 of which 38 were unique. While the automatic metadata generator produced 25 tags of which 24 were unique.

The students' results produced a power law curve, and is similar in nature to those described by Halpin et. al [40]. We then integrated across these values to separate the head from the tail (shown in Figure 3.1, for our discussion the assumption was made that the tail would contain 50% of the total tags provided). When overlaying the tags produced by the automatic metadata generator and the instructors we noticed several interesting points:

[22] To normalize the tag sets we did the following: we removed all extraneous punctuation, and converted all words to lower case. If the set included commas or semicolons, we created tag phrases out of each delimited item that was more than one word long, and individual tags out of those items that were only one word long. For each set that included no commas or semicolons we considered each word to be a tag and did not create any tag phrases. Finally, all words were stemmed using the Porter Stemmer provided by the Natural Language Toolkit [http://nltk.sourceforge.net/].

- 39% of the tags provided by the metadata extractor were within the head, while only 28% of the tags provided were in the tail (the remaining 32% were tags that students had not suggested).
- Hits within both the head and the tail were less for the experts, coming in at 25% for each. Half of the tags provided by the instructors were new tags. In addition, out of the tags that instructors agreed on, 66% of these (four of six tags) were not in the set of tags the students provided.

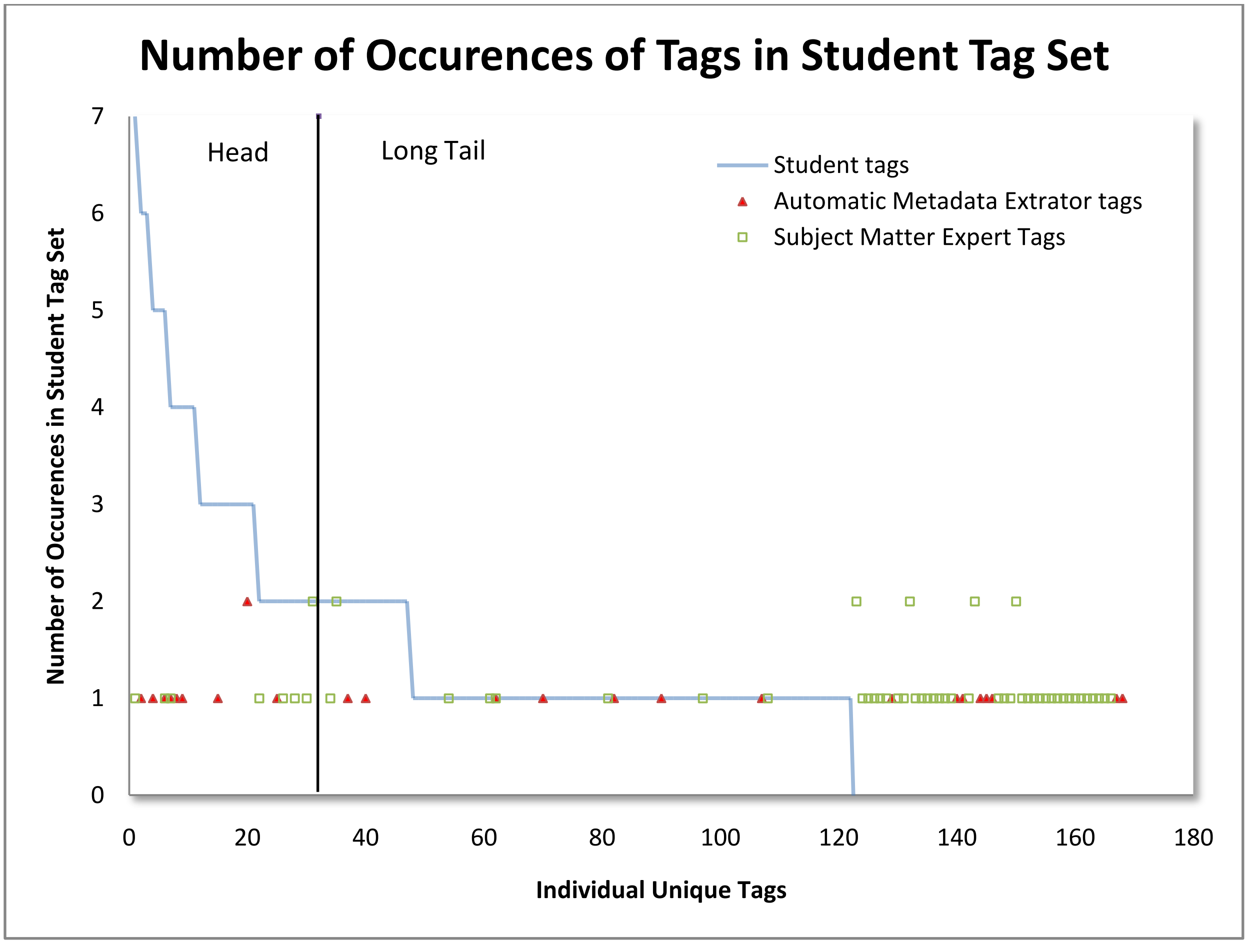


Figure 3.1 A plot of the occurrences of tags collected in the tagging pilot study, against those collected from different sources.

3.2.4 Discussion

Our initial tagging study has given us several new questions with regards to the implications of tagging in e-learning. Given that 50% of the expert tags were tags not within the body of tags used by students (head or tail), we question the benefits of providing these tags to students at all. The lack of expert time and willingness to fill in metadata has been cited [32][49] as a significant hurdle to deploying learning objects. If expert tags provide limited value to students, it may be more appropriate to bootstrap data sets with automatic tagging features and reduce the load on those who are creating content. We note the potential pedagogical benefits of collaborative tagging as suggested by [50]: that the tags themselves represent the expertise of the users. This suggests that at a collaborative level, a tag set can be inspected as the course is being given by the experts to gain an insight into the topics and concepts that learners are distilling from the online material.

Beyond the issue of expert time is the issue of control in the classroom. Unlike the open web, where individual success is evaluated by the individual, success in e-learning systems is typically dictated through a series of educator prepared exams. We have observed that educators are hesitant to change their teaching to adopt new methods in the classroom (virtual or otherwise), because of a loss of control. By engaging educators actively in the process of creating tags, it may reduce their fears of these new technologies. However, our results showed only 50% of the expert tags were represented in the tags of the students and, of these, only half were in the head (perhaps perceived as the most important tags by educators). Thus educators would need to be convinced that the long tail of tagging is also important. We also would like to suggest that unlike open an web system, the educator in the classroom is not merely a peer, and their tags may be more relevant to the examinations, which may be useful to learners. Therefore end-use of tags in an educational context is of significant interest to us.

We have observed that there is a good level of agreement between the automatic metadata extractor and students (a 67% occurrence of the former in the latter). This was surprising: we assumed that the learners would have a broader range of context (e.g. activities in the course, previously learnt materials, etc.) and would use this to alter their tag sets. However, it must be noted that perhaps there was an influence from the phrasing of the task for students. They were told to add keywords as if they were searching for the page in Google. As such it could be that the metadata extractor finds keywords similar to those found in a search engine and the students were attuned to creating keywords that generate desired outcomes using search engines. However, it is difficult to speculate on how tags and search engine terms differ. On the one hand the goal is for personal recall, and the other it is to use those which you expect to return documents of interest; the main difference would arise from how well one would know the domain of search and thus can incorporate tags or search terms not directly in the document. We contend that since the students were novices to computer science (the class was Introduction to Computer Science for non-majors) that they did not know the domain beyond the extent of what was covered in the learning materials and we feel that keywords and tags would have high co-occurrence since both keywords and tags would have been limited to those key concepts in the pages, when excluding self-reminder tags (e.g. the tag “read this” or “interesting”).

We are interested in applying the method we described here, along with several other cluster-based groups (e.g. [19] [43]) to see if co-occurrence can be strengthened. If so, automatic metadata extraction (extraction of keywords) may be an appropriate seed for educational communities (though we note that there should be benefits of the long tail as well, and the effects of such seeding would have to be studied to ensure a healthy tagbase remains).

In particular we know that there would be some disadvantage to users in new social software communities, as they miss the benefits of indirect SNS. Through, our assessment we have compared tags from other sources and we feel they offer some possibilities for new communities to overcome cold start issues.

3.3 **Summary**

We have discussed particular intuitions about how collaborative tagging applied in e-learning has interesting implications. We then turned our focus to a pilot study to investigate logical first questions about situating tagging in e-learning. Tagging systems, like most forms of social networking software, require a critical mass before they provide support to a community. In an online class, since learning content is constantly evolving over subsequent offerings of a course, it is most likely that the contributions in social software would have to start from scratch every time the class is offered. We are interested in pursuing this further, and quantifying the effects of applying various seeding algorithms against the growth, sustainability, and satisfaction of a learner community. Can we leverage automatic extraction methods (data mining) to overcome the cold-start issue with collaborative tagging in e-learning? Would such results be applicable to any social software community as it is confronted with a cold-start?

Chapter 4

The Open Annotation and Tagging System

This chapter[23] describes the Open Annotation and Tagging System (OATS), which was developed to investigate a number of social annotation techniques for current e-learning systems; these social techniques include collaborative highlighting, note-taking and tagging. While there has been much research on web annotation systems, none provides a combination of collaborative annotations, with a modular architecture designed for simple integration in learning management systems (LMS). This chapter is divided into two logical parts; the first part describes the features and the architecture of OATS in terms of the previous research, while the second part describes the results and lessons from an exploratory case study deployed for discussion around readings in a computer science social issues course.

OATS draws on the work started with a previous system, AnnotatEd, at the University of Pittsburgh. AnnotatEd allows students to highlight text in HTML learning objects and describe the learning object with descriptive notes [28]. Like AnnotatEd, OATS aims to solve a problem with the growing breadth of information made available in learning management systems (LMS). OATS is currently a fully functional modular open-source system, which extends and augments the highlighting metaphor by providing a collaborative tagging interface,

[23] This chapter up to and including section 4.6 is an updated version [5].

among other features. The aim is to increase the integration process for systems based on the latest Web scripting technologies, and thus encourage practical applications to be developed based on these technologies.

4.1 **Related Work**

4.1.1 Web Annotation Systems

The growing research into web annotation tools can be classified in two groups. The annotation systems of the first group focus on developing an advanced architecture and building a more sophisticated but simple-to-use annotation interface. Such systems try to support content-aware annotation on any arbitrary web page. Gibeo [73] is an example of an interesting web annotation system. Once the user is registered with the Gibeo website, in order to annotate any random web page, the user merely needs to add ".gibeo.net" to the URL. When any part of the text on the page is highlighted, a set of options is displayed to allow the user to specify the quality of the highlighted text, with labels such as "important", "wrong", or "cool". The users can also add comments, corrections, links, or shared discussion to any part of the text. Every annotation is shared with all users of the system and clicking on the annotation provides detailed information, such as the annotation author. Marginalia is a Javascript web annotation system that focuses on providing intuitive functionality for any arbitrary web page as well as the Moodle LMS [85]. Marginalia allows users to highlight any part of the text and write associated comments in the margin of the pages. Annotations may be marked as public or private.

The annotation tools of the second group use annotations for collaboration over the web. In contrast to the first group which focuses on the best way of collecting user-information for the individual, the second group instead aims to improve the sharing of information. The most famous tool in this group is Annotea, which enhances the collaborative development of the Semantic Web via shared web annotations. Annotations are in the form of comments, notes, explanations, or any other type of external remarks attached to any web document or portion of the document. The users are able to access all attached annotations when they open the original document [52]. Another popular tool is a "web discussion" feature in Microsoft Office 2000 that allows collaborative annotation of any web page. Cadiz, et. al. [18] studied the application of this annotation tool in the collaborative writing of a large product group at Microsoft and reported quite a variation in the usage of these annotations.

4.1.2 Social Navigation Support Systems

The AnnotatEd system (discussed in section 4.1.3) provides two types of SNS[24]: traffic-based and annotation-based. *Traffic-based* SNS relies on the traditional footprints concept in social navigation [82]. It generally provides information about the number of visits users have made to each link. Traffic-based SNS promotes links, which have a higher number of visits. For example, Dieberger and Lonnqvist [24] modified the collaborative web known as the CoWeb to visualize traffic-based social navigation. The system tracks how often a page is accessed or modified. It visualizes the density of the aggregated access for the past 24 hours by applying three levels of color intensity to the footprint symbol. *Annotation-based* SNS provides stronger support by employing the annotation activities of the users instead of the number of visitors.

---

[24] This section briefly surveys social navigation (SNS) support systems in terms of those related to OATS, for a description of SNS see section 2.1.1.

These systems promote links to pages annotated by users, especially pages with higher numbers of annotations or positive annotations. For example, Kurhila et. al [54] provide an annotation tool which allows learners to associate comments with a document. It provides simple annotation-based SNS by keeping track of when comments are modified and visually demarcates new comments.

CoREAD is a system that focuses on shared highlights in digital reading materials. As new readers come to a page in the system, users are able to see the highlights made by other users. Co-read uses the concept of annotation decay, as the highlights of previous users become faded over time. A small-scale study of the CoREAD system provided evidence that highlights would stabilize relatively quickly (after a few users) representing consensus of the highlighted text and be sustained over time [20]. This may be a similar process to that of tag stabilization (as described in section 1.2.2).

4.1.3 AnnotatEd

AnnotatEd (Annotations for Education) provides social navigation and annotation services for browsing collections of linked web educational resources. AnnotatEd was originally implemented to support the Knowledge Sea system [29]. Knowledge Sea is designed to help students navigate from weekly lectures to relevant online educational materials, in a map-based, horizontal navigation format. AnnotatEd has been incorporated in Knowledge Sea to provide annotations and traffic-based SNS. However, AnnotatEd is independent of the Knowledge Sea system and can be used with any linked collection of web resources. AnnotatEd accompanies the learners from page to page by redirecting all the links inside the page through the AnnotatEd server.

AnnotatEd allows students to annotate web pages by placing free-format comments that will be associated with the whole page or by highlighting specific parts of the page. In addition to allowing learners to annotate web pages, AnnotatEd also augments links inside the pages by measuring the past activity of the group in order to offer social navigation support. In addition to indicating group activity, link annotations also indicate individual activity of the user. Each link inside the page is annotated with two different icons shown in Figure 4.1.

Figure 4.1 Navigational Cues in AnnotatEd

The Traffic icon represents traffic-based SNS. The background color represents the magnitude of visiting activity done by the group of users. A darker color represents a higher number of visits. The human figure represents individual visiting activity. The human icon is presented with different shades of blue and the density of the color represents the frequency of the individual's visits. The *annotation* icon represents annotation-based SNS. Similar to the *traffic* icon, the background color represents the amount of group annotation activity and the color of the sticky note icon represents the frequency of individual annotation activity.

4.2 **The OATS System**

The OATS system provides a set of tools inspired by related systems that have emerged from work in web annotations, social navigation, web service architectures, and e-learning. So, while many of the individual functionalities are not unique, their combined use (including the modular system architecture) is unique, and is an extension of the work that has been started with the AnnotatEd system. In this first version, OATS is currently limited to annotations on text, as most of our learning objects are text-based HTML pages; although we have implemented prototypes for image and video annotations (as described in [6]).

The first goal of OATS is studying the benefits and problems associated with social annotation tools for students and instructors in a LMS. More concretely it is to provide tools to effectively organize and navigate learning content. In doing so, the learners make available information that could benefit their classmates by making available their personal perspectives to peers.

The second goal is to provide a proof-of-concept of how to decouple value-added services from a LMS, while providing all of the functionality of a tightly integrated solution. This is an important focus, as the achievements in the field of intelligent and adaptive e-learning systems have been significant, but they have yet to become main stream. This is because the applications created in research often focus solely on a limited set of features to benefit users. Thus, the cost of integration of a “closed” research system is too high. Methods that decouple functionalities and content seek to overcome this hurdle [15].

4.3 **The OATS Interface: Creating Annotations and Tags**

4.3.1 Accessing Functionality

To create an annotation in OATS the user simply starts by highlighting any piece of text in the content, by performing a click-and-drag. This will highlight the text, essentially changing the background color of the text in the document object model (DOM) of the learning object.

Whenever the user returns to that particular page, the highlighted text she has previously made will re-appear for her, as shown in Figure 4.2.

Figure 4.2 OATS displaying highlights within the iHelp LMS

OATS provides two methods for accessing functionality. The first method is through a drop down menu which positions itself in the top left hand corner of the learning object. This drop down menu provides access to the system on/off toggle, to user highlights and group highlights, and to a search/browse function to find any annotations in the system.

Figure 4.3 shows the expanded menu which appears when the user mouses over the unexpanded menu.

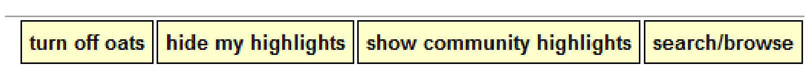


Figure 4.3 The OATS menu expanded.

The second method for accessing functionality is by clicking on an existing piece of highlighted text. The functionality available by clicking directly on a highlight includes searching and categorization, as well as the ability to add notes and tags, and to delete the entire highlight itself (see Figure 4.4). Deleting the highlight will also remove any associated tags or notes automatically. Users may also click on group highlights, but may only read the highlight's associated information. The interface provides a link for automatic inclusion of the highlight, which would allow users to add tags and notes.

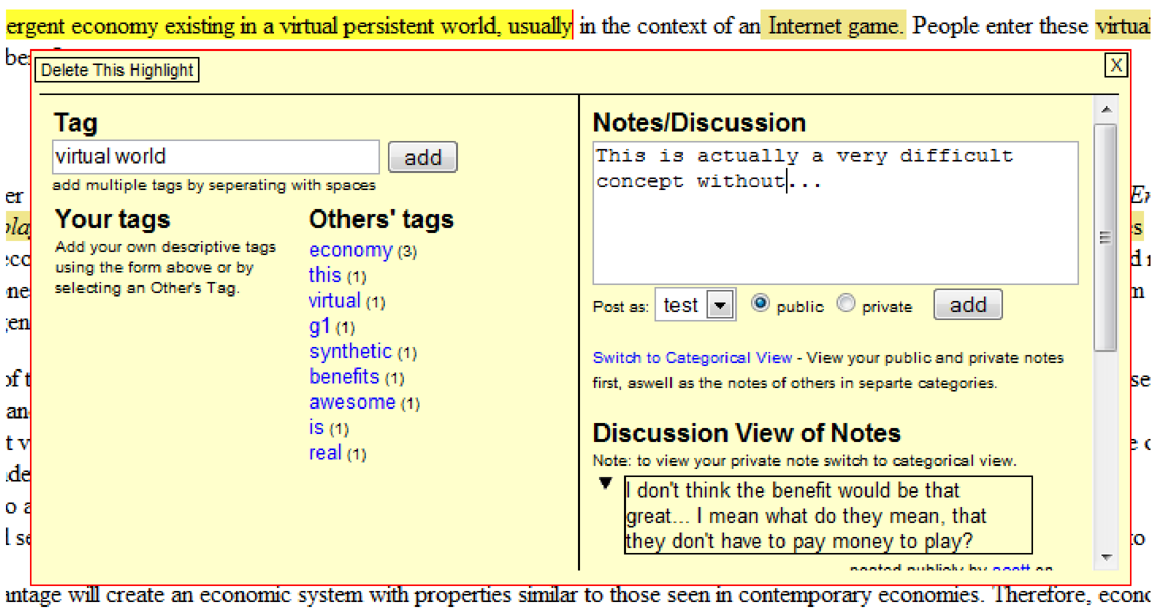


Figure 4.4 The OATS popup, which appears from clicking a highlighted piece of content. The left hand pane of the popup shows the interface for tagging. The right hand pane contains the interface for adding notes.

The entire interface is based on a simulated popup, which is styled to look like a sticky note. A simulated popup is implemented instead of an actual popup to avoid issues with popup blockers, which are now integrated into modern browsers to block unwanted advertisements and sites from appearing. Instead the highlight system writes the popup to the page in a hidden iFrame HTML object, which positions itself and becomes visible on appropriate events, such as clicking on a highlight.

The rest of this section will discuss the actual process of adding and maintaining tags and annotations. Section 4.4.1 will discuss what aspects of the system encourage learners to tag and annotate and what benefits they will receive.

4.3.2 Tagging

Tags in the OATS application are a free piece of text which contains no spaces, as they are with most tagging systems. To add tags the user first must click on a piece of text she has already highlighted. This provides her with a popup, which has been contextualized to the highlight that has been clicked. Finally, the user may add many tags by listing them in the text box and separating them by spaces. When a tag is added, it appears under the user's "My Tags" area, while tags added by any classmates appear under "Others' Tags." The other tags are determined by providing the most popular tags added to any highlights of other users which share any overlap with the current user's selected highlight.

4.3.3 Note-Taking

Note-taking in OATS refers strictly to adding longer pieces of free text. When compared to tags, notes have absolutely no rules, and are used for the purpose of adding normally structured sentence-based messages.

To add a note the user first must access a highlight-contextualized menu by clicking on an existing highlight. To add a note she simply types a message in the text area. She has the added capability of making notes "public" (globally viewable) or "private" (viewable only by her); the default is global which can be changed by checking the "Private" checkbox.

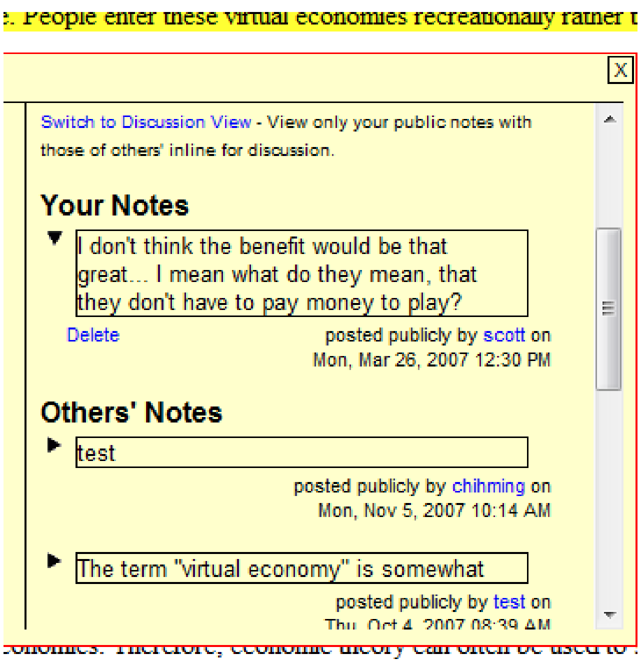


Figure 4.5 The alternate Categorical View for notes.

There are two different options for viewing notes. The first is the "Discussion View of Notes", which lists all notes associated with a region of text by listing them chronologically. This allows users to use the notes as a discussion forum (without threads), where they can ask questions, respond to questions, or post relevant links. This view does not show the user's private notes (this view can be seen in Figure 4.4). The second view is the "Categorical View" which can be accessed though a link below the note entry field. This view separates the current user's notes from those of others. In this view, the user's notes after being added will appear under

"Our Notes." Any notes added by other users about this highlight will appear under the "Community Notes". All notes contain information on when and who added the note (see Figure 4.5). A user can delete her own notes by clicking the "Delete note" link which is located under each of her own notes.

4.4 **The OATS Interface**

4.4.1 Motivating Participation

Value-added services, refers to the set of functionalities which are added to the core functionalities of an LMS. In the case of OATS it refers to the ability to add highlights, tags, and notes. Further, it entails the added benefit of additional support that learners can receive when the influence of the overall user community is made available.

OATS was designed with the goal of providing enough motivation for users to highlight, tag and annotate text without the added benefit of community support. While, we think the community support is an important selling point of the system, the functionality at the individual level must sell itself. This is in part due to the issue of the "cold-start", where the first users to highlight in the system will not have any benefit of the information provided by others in the community. It should be noted also that the community benefits are therefore considered to be provided through implicitly gathered information, as the community based information is a consequence of users acting primarily for individual purposes.

4.4.2 Tag Browsing

Tag Browsing - in terms of an individual user - is an interface to automatically categorize information based on tags. In the community sense it is a way to gain a "global view" of the tagging of the entire community, while still allowing users to drill-down on the individual contributions of peers. In this sense the tag browsing interface is based on pivot browsing (described in 2.1). This functionality can be accessed through the drop down menu, and provides three options: “My Tags”, “User Community”, and “Community Tags”.

Information provided by the individual user is located under “My Tags” in the interface. By clicking on “My Tags” a list of all the tags the user has used is revealed, which are ordered from the most to least frequently used tag. By clicking on each individual tag a list of three options pertaining to the tag are presented: “Search for *this tag*”, which links the user to the search interface; “View pages you’ve tagged with *this tag*”, which shows all pages that contains highlights which have been described with the given tag; and “View highlights you’ve tagged with *this tag*”, which presents the highlighted text that has been described with the given tag. An example of these functionalities is shown in Figure 4.6. Links are provided to the pages so that the highlights can be seen in context. The community aspect of the Tag Categories is shown in

Figure 4.7 and Figure 4.8. By expanding the “User Community” section of the browsing interface, the user may access a list of all contributing users, ranked by the number of highlights they have made to date. Each of the users’ contributions are displayed by type (the total number of highlights, Tags and Public notes). Expanding a user section will give access to the selected user’s “My Tags” interface, where every user’s public contribution can be explored in depth.

By expanding the “Community Tags” section (see Figure 4.8), the most popular tags are shown in descending order of number of times used. This gives the user an idea, at a much higher level, the overall view of all the content. The user can get a sense of what are the most important terms and/or ideas at a course level. By clicking on one of the tags the user can select to add the tag to the highlight in context, or to search by the given tag.

Figure 4.6. Viewing "My Tags" in the Tag Categories interface.

Figure 4.7 Pivot browsing based on user and tags. Users are ranked by the number of highlights they have made, and the total contributions of each of the users is displayed. Expanding a user section will give access to the selected user's "My Tags" interface.

Figure 4.8 Viewing "Community Tags" in the Tag Categories interface.

4.4.3 Search

Searching for tags, notes and pages is an important aspect of the system. Along with tag browsing, the searching functionality would potentially provide the user with more motivation to contribute to use the system.

Figure 4.9 shows searching for a page by tags used in the highlights of the page. The results are returned ordered by "score". The score simply counts all occurrences of the tag in any user's highlights on the given page. The page with the highest count of tag occurrences is presented first.

Searching for notes provides the actual notes that match the search criteria as shown in Figure 4.10. This allows the user to quickly find relevant user augmented information, and to ascertain to which page it pertains. When the tag searching option is used for searching tags, notes that share highlights with the search term tags are retrieved.

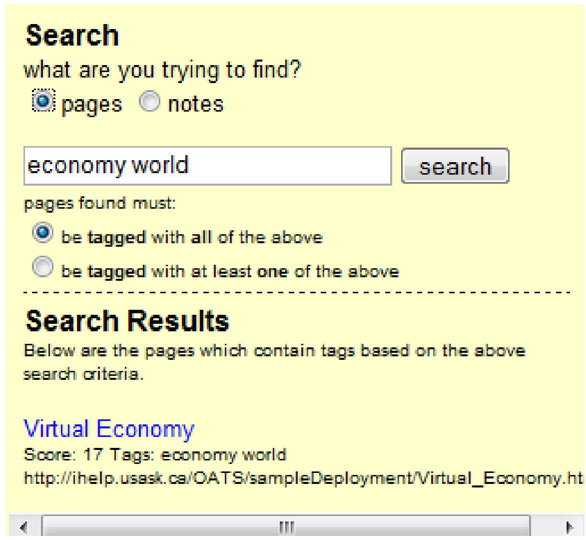


Figure 4.9 Searching for pages by tag.

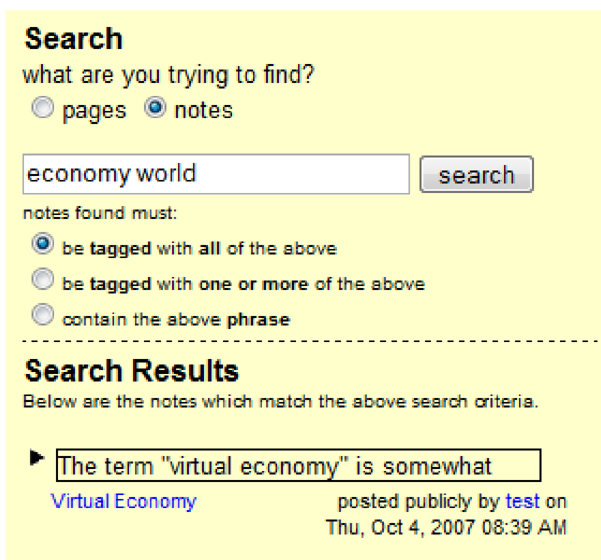


Figure 4.10 Searching for notes based on highlights tagged with the search tags.

4.4.4 Community Highlights

Hill and Hollan created the idea of “read and edit wear”, which provides a visualization of users’ read and write actions in documents. The read and write actions are displayed via a histogram in the scroll bar of a programming environment. This allows rapid navigation to the parts of the code found to be most important by other users [45] [46]. We provide a similar functionality based on the notion of “wear” and similar to the CoREAD system (previously discussed) for providing a text-inline visualization of highlights created by users. This allows users to quickly see what others thought were note-worthy parts of the text.

By clicking the “show community highlights” button in the drop-down menu, the visualization is enabled. The visualization works by calculating the overlap of highlights in the text by users other than the current user. We currently have set four levels that show the concentration of highlights: none (0 highlights), low (1-3 highlights), medium (4-8 highlights), or high (9 or more highlights). These translate into different thicknesses of highlights within the text, shown in Figure 4.11. For instance, in the case of no support the text remains unchanged, for low to high the pink line becomes thicker as the concentration of highlights increase. This particular functionality relies exclusively on other users having contributed by highlighting text on the particular page, although having no contributions could also give the user some useful information. Such social information could indicate the importance of the content or if the user is ahead of other students in consuming the content.

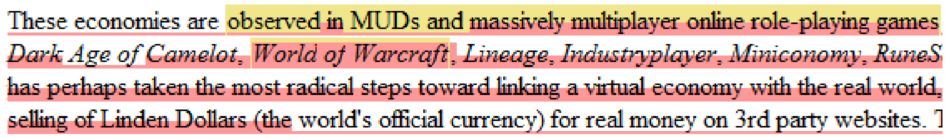

These economies are observed in MUDs and massively multiplayer online role-playing games
Dark Age of Camelot, *World of Warcraft*, *Lineage*, *Industryplayer*, *Miniconomy*, *RuneS*
has perhaps taken the most radical steps toward linking a virtual economy with the real world,
selling of Linden Dollars (the world's official currency) for real money on 3rd party websites. T

Figure 4.11 A visualization of community support based on highlighted text. The Yellow text represents highlights suggested by the current user. Pink highlights represent the 3 levels of selected text by other users. "World of Warcraft" has been selected by the current user and also shows the highest concentration of selection by other users.

Some research needs to go into exactly what is the best way to calculate these different levels of community support. We hypothesize that user's will find any reasonable algorithm useful, then the case without any visualization. However, it may be that some ratio of highlighted text on the page, to discourage what users would see as an empty system. For instance, in the case of a "fresh" system or in a class with a smaller user base the first users would see no contributions from others, and the concentration of highlights would fill up very slowly. With such a ratio-based visualization, the system would seem to fill up with annotations more quickly.

The concentration of highlights on the page, indicate the community's support of the particular text being note-worthy on the page. On tagging sites, like del.icio.us, using a tag is a vote for it being recommended to others for the particular resource. This has been extended in OATS to a part of the content in the page, where the user can simply read a histogram embedded in the text to get an idea of the important content.

4.5 **OATS System Architecture**

One of the main goals of OATS is to give a complete set of value-added tools that can be used in systems both at the University of Pittsburgh and the University of Saskatchewan, independent of the LMS in which they are deployed. This is closely related to some past work on system architectures that is aiming to decouple different components from adaptive e-learning environments as described in the Adapt2 protocol [15]. The Adapt2 protocol mainly focuses on the decoupling intelligent content (that is content that creates and makes use of its own sophisticated user model, such as interactive simulations or activities) from the LMS. The OATS system is the first approach that we know of which seeks to completely decouple value-added functionality from a LMS or learning portal. Other "modular" LMSs such as Moodle[25], still rely on integration of modules based on a system dependent API.

To achieve the decoupling of functionality from the LMS, the OATS system was divided into two parts. The first part is the OATS Webservice implemented as a Java servlet. The webservice manages and maintains all of the logic and stores information on the users and their tags and annotations. The second part is the OATS Client which is written in Javascript. It is responsible for the user interface and delivering information between the user and the OATS Webservice.

4.5.1 OATS Webservice

The OATS webservice interface is a Java servlet, using a MySQL database and is implemented as RESTful webservice. Representational State Transfer (REST), refers to a design pattern for distributed hypermedia systems. A full discussion of REST is outside the scope of this thesis; instead we refer the reader to Roy Fielding's thesis [30]. While, we do not argue for REST style web system interfaces for all systems, it dictates a style of accessing web based

[25] See http://moodle.org

resources, without the communication layer needed with webservice protocols like SOAP (Service Oriented Architecture Protocol). This is achieved by using a naming pattern for resources and existing request methods provided by HTTP. We feel because of the reduction in messaging overhead it may be particularly well suited for decoupling functionality in systems such as a LMS, where a number of different client libraries may need to be integrated each corresponding to different functionalities.

For example, consider the following scenario: a “GET” request to the OATS Webservice URL http://oatsserver.org/tags/user/johndoe would return an XML document containing all the tags that the user with username “johndoe” has used. While adding a tag would require a “POST” request to the same URL with message body “url=http://lmsserver.com/lesson1/aboutHCI.html&tag=hci”. In this case the user with username “johndoe” has tagged the page at “http://lmsserver.com/lesson1/aboutHCI.html” with the tag “hci”.

4.5.2 OATS Client

The OATS Client is a Javascript library which is included in an HTML learning object. The inclusion of the client library can be done in two ways: either by manually including it using an HTML “script” tag, or by having the LMS write the Javascript files to the learning object’s document object model (DOM). Once included the client library writes the OATS interface to the DOM of the learning object, and communicates with the OATS Webservice to obtain any annotations to be initially displayed on the page for the present user.

The client communicates with the service through Asynchronous Javascript and XML (AJAX). This, coupled with the manipulation of the DOM to display highlights and the interface, allows OATS to be integrated without disturbing other unrelated functionality and for the system to work "out of the box".

It should be noted that one significant limitation of such a system is the security requirements of the browser to access the XML over HTTP request object used by AJAX in modern browsers. To avoid the security problems associated with cross-site scripting, the AJAX library requires that the Javascript source and the URL of communication requests be located on the same server and port as the web page in which they are embedded. With regards to the Javascript source, we don't view this as an issue because the text-based Javascript libraries are small enough not to cause any conflicts or problems, and should be easily hosted on the same server as any web-based LMS. However, the requirement for the OATS Webservice (which requires a servlet enabled webserver and a MySQL server) to be located on the same server as the LMS is too great for the claim that OATS is truly decoupled. We provide a simple solution to this problem which is to provide a proxy on the LMS server that simply forwards communications directed from the client to the webservice and back, as suggested by Crane et. al. [23].

We have implemented a proxy in Perl/CGI to prove this concept in a research project which was based on integration with the Knowledge Sea system at the University of Pittsburgh, and carried out as a study at the University College of Dublin, Ireland in November 2006. The result was that the OATS client library was integrated into the Knowledge Sea portal from the Pittsburgh servers. The Dublin participants then had a local OATS client by which they communicated with the Pittsburgh servers. The proxy written in Perl/CGI[26] forwarded requests and received responses with the OATS server which was located at the University of Saskatchewan.

4.5.3 Integrating OATS In Practice

Based on its *open architecture*, OATS it is easily integrated into learning content regardless of the LMS in literally minutes, by including several Javascript files on the server. OATS uses an external database which is completely independent of any knowledge of the LMS. Thus, for example, including the Javascript source files in course content, which can be uploaded as a content package to the iHelp LMS.

The only further possible requirement for integration is the availability of a unique user identifier (id). A user id is needed by OATS from the client LMS to distinguish between each user. By default OATS looks for a "userid" parameter in the URL of the page in which it is integrated or in a parent frame. However, this can be easily changed to allow for a user id in any element of any related document or window of the LMS.

It should also be noted that OATS can be used independently of any LMS, and thus be used as a more general purpose web annotation system outside the domain of e-learning.

4.6 **OATS USE CASE STUDY**

An exploratory case study of the OATS system, was carried out in the Computer Science 408 class (Ethics and Computer Science) at the University of Saskatchewan, the during

[26] The proxy could have been written in any language with socket communication capabilities.

winter semester in 2007. The class was a suitable place to evaluate the OATS system, since part of the student's evaluation was based on participating in on-line discussions. The class evaluation was based upon the reading and discussion of class-relevant articles from around the web, which were posted each week. This work was completed outside of class time, and performed when convenient (given a final deadline, the end of the semester) and using the computer of the student's choice. Firefox version 2.0 or later was required to use OATS (although the system works with Internet Explorer 7, there are some small presentation bugs). All students had access to the Firefox web browser.

During the first 8 weeks of the course, the articles were discussed using the iHelp Discussions discussion forum. In the final two weeks of the course OATS was used instead of iHelp Discussions for posting discussions and sharing links. iHelp Discussion is a "typical" web discussion forum and can be compared in functionality to other more widely used web-based forums such phpBB[27]. iHelp Discussion has had many iterations of design and development, and has been widely used by Computer Science students for dozens of classes, and thousands of students over several years. To this end the iHelp Discussions application was used as a gold standard against which to gauge students' impression of OATS in terms of usability and functionality, since it is both typical of more traditional forums and was used by all students within the context of the course.

[27] See http://www.phpbb.com

4.6.1 Case Study Task

The task was for students to use the note-taking functionality for discussions of 3 articles over the final two weeks of classes. The main difference between the case study discussion task and the discussion task previously completed in the course was that OATS allowed students to seed discussion directly in the reading content since notes are applied to highlights, rather than using different systems for reading and discussion as in iHelp. To this end we expected students to enjoy the time-savings of using a single system where they did not need to change contexts to make comments. Further, because these notes and highlights could be organized using tags, we expected that students would find tags a useful means to organize discussion-based information and would use them for studying for the course.

While participation in the study was not mandatory, students were required to participate in discussions using OATS, as discussion was mandatory to receive class marks (as was the case using iHelp). Students were assured that if they chose not to take part that none of their exemplary data would be used in reporting on the system. Regardless of their participation, students understood that data was to be anonymized for any use whatsoever. However, their identities would be known between classmates, to the instructor, to the experimenter and to the teaching assistant. This information was outlined both on several occasions in class to the students and in writing via an online consent form.

4.6.2 Case Study Methods

An in-class tutorial was given to introduce OATS 2 weeks prior to the system's deployment. This introductory tutorial simply introduced the functionality of OATS, the ethics implications and to explain the task. Care was taken not to influence the students in exactly how to use the system to achieve various pedagogical goals, but rather to simply show what was possible. The following week a webpage was posted, which was a summary of the in-class tutorial that had been given the previous week. The page covered how OATS worked, and also allowed the use of OATS to annotate the document itself and ask questions (please see Appendix A for the Introductory Webpage). Students were informed that this would be the only page where the researcher would interact with them using the OATS system. The experimenter was available to be contacted by email at any point during the study.

The online consent form was presented to students when they logged into the iHelp LMS, after the start date of the study. This was followed by a page based on the in-class tutorial page for OATS, and the papers to be discussed. The experimenter and two teaching assistants seeded the tutorial page with example highlights, which they categorized using tags and discussed using the note-taking functionalities.

The first week's readings presented two papers. The first was Thomas Friedman's New York Times Magazine article, "It's a Flat World, After All" [31], which was split over four separate HTML pages. This first article was also seeded, by the two teaching assistants, to give all students the benefit of seeing discussion and highlights for their first experience using the system. The second article was "The Long Tail" [1], which was split over five pages. The two articles totaled 10,465 words in length. The splitting of the articles was significant since the OATS prototype does not contain a preferences feature, so all settings would reset to default when loading a new page. This meant that if a user would prefer to have the "community highlights" function activated (see section Community Highlights) all the time, it would reset every time. Because of this and some feedback received during the case study, the second week's reading was switched to a single page format. The reading was “Five Things We Need to Know About Technological Change”, a transcript of a speech given by Neil Postman [70], which was 3726 words in length, and displayed on a single page.

As an external feature to the system a tag cloud was created on a page separate from the reading tasks. The tag cloud presented all tags that were in the system (at that moment), and mapped the frequency of a tag’s use to font size. The idea was to encourage tagging by students through showing an artifact which visually showed their tagging contributions. No data was collected about the tag cloud other than the number times the cloud was viewed.

Following the class, an email was sent to all consenting students to request their completion of an online exit survey. They were not required to fill out the survey, but were asked to do so in order to help the researcher and improve the system (please see Appendix B for the complete online survey). The survey was composed of a number of questions relating to previous experience with similar systems.

4.6.3 Case Study Participants

Of the 15 students enrolled in the class 12 consented to participate in the study. All students were upper year computer science majors. Of the 12 consenting participants we received 10 completed exit surveys. Two of the survey respondents were third-year students, while the other 8 were fourth-year students (see Table 4.1).

User Details	Totals
Total Users	15
Consenting Participants	12
Surveys Completed	10
Study for the exam?	5 studied, 1 did not study, 6 completed the task just before exam time

Table 4.1 Overview of the participants in the study.

Methods of organization in physical materials	
Highlighting	9
Marginalia	5
Writing keywords or concept words	7
Writing summaries	7
Use of other systems with OATS features	
Highlighting, systems with	2
Tagging, systems with	2

Table 4.2 Participants' reported experience with schemes of organization (physical or software). The number of positive responses of the 10 participants is given.

Table 4.2 shows the participants' reported previous experience with physical methods of organization and software systems with aspects similar to OATS. For each case the number of participants of the 10 consenting users is reported. Of note was the participants' wide-spread experience with using both highlighting (9 of 10) and writing concept keywords (7 of 10) as means for organizing physical materials. This did not translate into use in software as only 2 of 10 participants had used either tagging or highlighting in software systems other than OATS.

4.6.4 Case Study Results

Results from the case study are separated into two types: usage data and communication data. Usage data was collected from the iHelp LMS database where users were viewing the readings, and from the OATS system database itself. We also selected usage data for comparison from the iHelp Discussion Forum, where users had previously been discussing class readings before the deployment of OATS. Communication data was collected through email communications with participants and with the teaching assistant who had sent comments and bug reports directly to the experimenter, and through the completed exit surveys (not including background experience data reported in the previous section).

4.6.4.1 Usage data

Table 4.3 shows a summary of all OATS usage data in the study by consenting participants. Several aspects of the data are of note:

- One user accounted for over 40% of all tags applied.
- All users created at least 5 highlights; however one user failed to make a note despite having this as part of the class evaluation.
- The most popular action was to turn on community highlights with each user activating this feature an average of 29 times. This is despite having a total of only 10 pages in the readings (meaning users activated community highlights an average of ~ 3 times/page).
- There was very sparse usage of search and browsing functionalities, in spite of the fact that 5 users returned to the system to review for the exam (this was assessed by the researcher viewing access patterns). 1 user did not return to the system to study, while 6 users used the system for the first time just before exam time.

System Action	(all users) Totals	Min	Max	(per user) Average	Std. Dev.
Highlight Created	137	5	29	11.42	7.50
Tag Applied	168	0	70	14.00	19.84
Note Created	88	0	16	7.33	4.96
of which are private Notes	9	0	9	0.75	2.60
Viewed Community Highlights	349	7	63	29.08	18.06
Viewed Tag Cloud	22	1	3	1.83	0.72
Browse Users	8	0	5	0.67	1.50
View Community Tags	8	0	3	0.67	1.07
Browse My Tags	10	0	5	0.83	1.75
Browse My Pages By Tag	5	0	5	0.42	1.44
Browse My Notes By Tag	5	0	2	0.42	0.79
Browse My Highlights By Tag	2	0	2	0.17	0.58
Browse [Other] User Tags	2	0	2	0.17	0.58
Browse My Pages By Tag	0				
Browse My Notes By Tag	0				
Browse My Highlights By Tag	0				
Search for Pages By Tag	9	0	4	0.75	1.36
Search for Notes By Tag	0				

Table 4.3 A summary of all OATS usage data in the study.

4.6.4.2 Communication data

In this section we will present the data collected from the exit survey along with noteworthy comments from email communications with participants about the system. The presentation of the data will be separated into two parts: responses to semantic-scale questions and noteworthy comments. We present the data collected from the semantic-scale questions in Figure 4.12. Our Semantic-scale questions were based around gathering students' thoughts on 3 main points: the suitability of OATS for the case study task in direct comparison with the iHelp Discussion system, the usefulness of the main interface features of OATS, and general impressions on the usability of OATS.

We present the semantic-scale data in a box plot. We will describe the visual features of the box-plot for the reader unfamiliar with this charting technique. Since, semantic-data from users provides a non-normal distribution of data it is best characterized, using measurements of the minimum, maximum, and percentile positions of the responses collected. Centrality of the data should, therefore, be illustrated with either mode or median, but not mean. The box plot is a nice method to visualize such data, since it provides all of the appropriate data concisely. The bottom and top of the yellow box, show the 25th and 75th percentile of the answers respectively. The thin dark lines extending to the small black dashes represent the limits of all answers, minimum and maximum. The grey dash shows the position of the 50th percentile. The mode of the responses collected is shown by the black diamond.

Figure 4.12 shows the response to each of the semantic-scale questions, where the x-axis shows each of the statements made, while the y-axis shows the level of agreement with the statement (please see Appendix B for the exact question phrasing). The scale used for level of agreement was a 7 point scale (score values could range between 0-6, please note the 7 on the vertical scale is an unnecessary artifact of the graph). We interpreted the scores mainly by considering the centrality of the responses, as indicated by the 50th percentile (effectively the median) and the mode. The size of the box and the extended hash marks indicate therefore the level of agreement with the central responses. The scale used for plotting the responses was:

0) Extremely unlikely/unuseful
1) Quite unlikely/unuseful
2) Slightly unlikely/unuseful
3) Neither (neutral)
4) Slightly likely/useful
5) Quite likely/useful
6) Extremely likely/useful

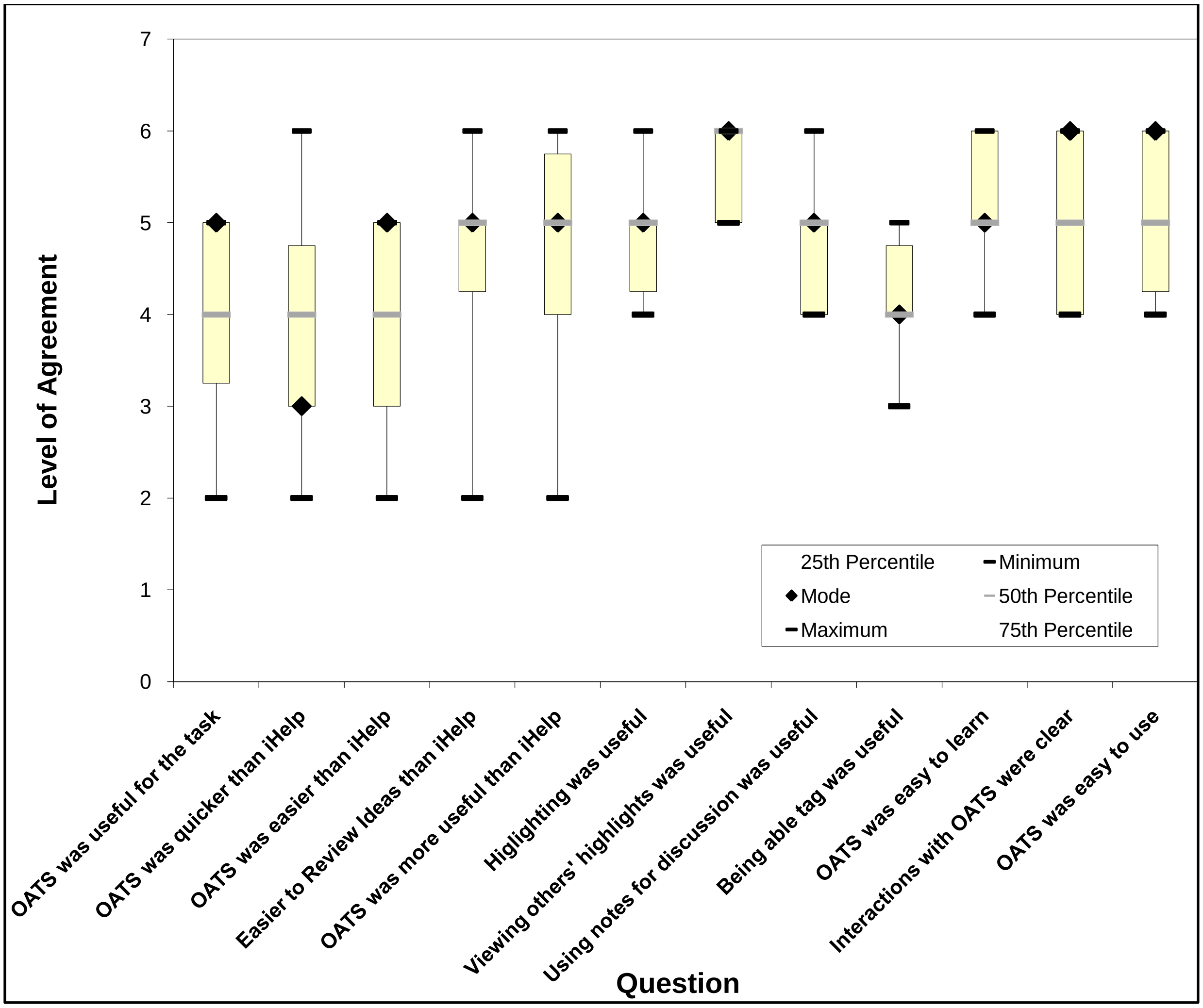


Figure 4.12 A box plot of results of semantic-scale responses from the OATS Case Study exit survey. The level of agreement scale ranged from 6 – strongly likely/useful, 3 – neither, and 0 – strongly unlikely/unuseful.

We point out a possible confound for the semantic-scale questions. This is due to the phrasing of the questions, and a possible effect which may have resulted from the desire of the students to please the experimenters. This could have resulted in a shifting up of the responses in the box plot. So, in interpreting these questions we are careful not to make broad conclusions, but rather we keep them to the context of the study and the task in which it was used. However, we consider the responses from the semantic-scale in combination with the free-form comments to form a more complete picture. Also, we consider the responses about OATS features (e.g the ability to highlight or being able to tag) as rankings of which parts are most useful, within the context of the study. With regards to the questions which compared OATS to iHelp, we interpret results that suggest OATS is superior to iHelp for the task in some aspect as meaning at least OATS is not significantly worse and is possibly better. We believe the effect of statement phrasing on participant responses would not be so strong as to cause them to give statements contrary to their actual beliefs.

We interpret the response from the box plot as one of three specific response types where 3 is a neutral response, less than 3 is negative response, and greater than 3 is a positive response. We consider those questions with a consistent central tendency (if not all) of the responses above neutral as being a generally positive response from the study group, but also look to strong evidence in the statements and actions of the participants to support a positive response conclusion due to our possible confound.

Below we provide significant or note-worthy free-form answers from the exit survey with a few comments taken from email communications[28]. We present comments around the questions: 1) What makes OATS useful? 2) What makes OATS not useful? 3) How can OATS be improved?

1) What makes OATS useful?

- "Going through an article to highlight and tag was very easy and enjoyable"
- "Helpful [*sic*] to have other peoples [*sic*] opinions, and suggestions linked directly to the material"
- "both highlighting, tagging are commonly used ways for readings"
- "When discussing an article on a forum, one has to quote their references (often with a link). It is very nice to be able to comment DIRECTLY on the text which pertains to the comment. Being able to see what other people think is important is also amazing. There are odd cases in which people highlight irrelevant text. However, I believe this is compensated well in being able to recognize lines which have been highlighted several times by users; as these thicker highlights are always important."
- "it acts as a big indexing of material we took. if other classmates highlighted something it helped me realize it's significance especially if i [*sic*] overlooked it"
- "It's easy to see where the interest has been concentrated which could give way to potential discussion. You don't have to read the whole paper in order to comment on a highlight, it is simple to become involved without spending hours ready hierarchical posts on i-Help."

[28] Permission to use quotes from email communications was received directly from the individuals.

- “The learning curve was minuscule and I prefer OATS to iHelp any day. … It is also more conducive to group interaction I think because people tend to rate their interest in a particular iHelp thread based on the title somebody gave it. The titles can be a bit misleading sometimes I find.”

2) What makes OATS not useful?

- “It is not a community chat/messaging system like ihelp”
- “not enough people participating in the discussions.”
- “Difficult to see whole discussion without clicking on all highlights (some of which are empty).”
- “I wasn't able to utilize the tag function as well as I wanted. I found myself adding a lot of tags, as I'm experienced with tagging, but very rarely searching for tags. Although, this may reflect upon the nature of the course; I think other more technical courses could offer a lot more of a benefit to tagging.”
- “… iHelp lends itself to a broader array of unrelated topics, whereas OATS tends to result in more focused discussion on a narrower set of topics. One could see this as good or bad I suppose, depending on the circumstance.”

3) Can you think of anything that would improve how easy OATS is to use?

- “The ability to be updated (via email, or internal integration) when posting is made would be an asset to this system. Also, allowing the system to be smart enough to recognize highlights that were new since last visit (with a new color maybe?) would be a significant improvement.”

- “It would also be nice to be able to track (with a different color?) the highlights of a specific user, such as the professor. Similarly, the ability to make specific user's highlights lighter could help "block" a user. There is possibility that a "karma" system could be implemented well, depending on the community using OATS. If users were able to rate each other, or their comments, users that had better "karma" ratings could have darker highlights, while users with low ratings could have lighter highlights. This may not be good for all situations, but could definitely help in a large community, specifically with adolescents.”
- “Perhaps some kind of familiar forum view that would perhaps hide the document but permit ease of viewing with quick referencing.”
- “I don't think that the searching utility was as good as it could be. More advanced search criteria could help this project quite a bit. Advanced tagging would also be a nice addition, for people who are familiar with tagging. An example of this is quoting words together ("World of Warcraft" rather than "World" "of" "Warcraft").”

4.6.5 Case Study Results Interpretation and Discussion

4.6.5.1 System Features

Highlighting

With regard to the functionalities or features of OATS, our data gathering particularly in the survey focused on three main aspects: highlighting, tagging and note-taking. Recall, the purpose of this thesis, and part of the larger motivations behind OATS, was investigating tagging and how it might best be applied in e-learning. This being said based on students’ actual use of OATS it would seem that the highlighting functionality was the most popular. All responding participants found that viewing the community highlights was either quite useful or extremely useful. Interestingly, viewing others' highlights outweighed the perceived usefulness of the participants creating their own highlights (most responses centered around quite useful). This is also supported by the usage data of OATS, where all users turned on the community highlights multiple times, 3 on average per page in the reading tasks. This is despite users needing to only do this once per visit to the page. This indicates that most likely several users returned to the reading multiple times to find new comments.

Users noted that highlighting was a familiar and natural task, and one even described it as "enjoyable". Group highlights seem to be a borderline case between direct and indirect SNS, since creating highlights is an explicit action. Creating a highlight is in effect saying directly to others, "This is important." A user noted that, "if other classmates highlighted something it helped me realize it's significance especially if i [*sic*] overlooked it." Since, the highlights are shared with everyone immediately they become almost a direct recommendation for attention.

Given the popularity and focus on highlights some of the suggested functionalities to improve the feature were quite insightful. In particular, users suggested schemes for weighting of highlights based on user reputation, indicators for unseen highlights using different colours, and support to make instructor contributions stand out. These are all very interesting possibilities focused in and around the idea of sharing highlights. We are interested in examining highlighting alone in depth; in particular we have not yet seen other work using our embedded histogram-like view for SNS in text. How to further enrich shared highlights, taking into account the viewing history of the user, along with the issue of how shared highlights emerge in these systems are interesting questions we intend to pursue.

Tags

As mentioned our initial research interest in the OATS system revolved very much around tagging and its application to e-learning. Overall, user responses centered on “slightly useful” for tagging given the task. However, we thought users would employ tags more widely. We had indicated in class that tags were used for organizing both notes and highlights in the system thought the search/browse interface, and so we expected they would be of use when reviewing for the final exam. Tags would allow them to more easily be able to find conversations and interesting text. While, there were more tags applied than highlights (168 tags to 137 highlights), over 40% of the tags used could be attributed to a single participant. These numbers further misrepresent the actual tagging practice, which can be seen from browsing the applied highlights in the readings. We could see in general most users started out by applying several tags to highlights, but then quickly abandoned the practice.

The most prolific tagger noted that, “I wasn't able to utilize the tag function as well as I wanted. I found myself adding a lot of tags, as I'm experienced with tagging, but very rarely searching for tags. Although, this may reflect upon the nature of the course; I think other more technical courses could offer a lot more of a benefit to tagging.” The background of students revealed that only 2 of the 10 respondents had experience with tagging. The prolific tagger seems to have noticed a major implication for tagging in the context of its application. The task of the case study was focused on readings and discussions, and so this did not necessarily mean students were going to study for the exam. In fact, our analysis of the system access suggests that only 6 of the 12 participants did study for the exam by accessing readings a second time. So, only half the participants would see any benefit from tagging and thus be motivated to do so. While the prolific tagger suggested that a more technical course might see more widespread benefits from tagging, we feel that there is a more important pre-requisite. Tagging, in most systems, is a persistent action, where tags last indefinitely (or at least for the life of the system in which they are used). In this case, tags were applied at the end of the course, in some cases only a day or two before the exam. We categorized 5 users as “late-comers”. These were users who came to the system to do readings and make discussion notes a week before the exam or less. These users had no real-motivation to tag (although some did, most likely simply to “try out” tags).

We believe since users took so well to highlighting, they relied on highlights to find contributions. There were only 10 pages in total in the task, so users found it easier to navigate by highlights and community highlights. In fact one of the most elaborate parts of the system, the browsing and searching interface, was accessed only a few times by all users. These interfaces relied on tags to allow users to interactively pivot-browse around tags and users, and to view discussion and highlights. It is not clear if the low access to the searching and browsing functionality was due to a weakness in the interface or to the simplicity of the task. We feel most likely it is the latter, although one user commented he would like to have “a more advanced” search interface.

Our findings suggest that for tagging to be applicable and widespread, it needs to be persistent and span some length of time, where recall without an organization scheme would suffer. This being said, we will look at the possibility of integrating OATS into the iHelp LMS, in hopes that users will more readily have a benefit of tagging, being able to apply them over entire courses or over several academic years (for further discussion please see section 6.2).

Notes/Discussion

Overall, users perceived the notes fields for discussion as useful; in fact most were centered on “quite useful”. However, based on the comments we received we feel the high level of responses is actually more related to the fact that notes were embedded in highlights, rather than the notes interface. Users felt very strongly about the ability to highlight, and tying further functionality such as tagging and discussion was also seen as a benefit. One user commented, “It is very nice to be able to comment DIRECTLY on the text which pertains to the comment.”, and another said, “[it is] Helpfull [*sic*] to have other peoples [*sic*] opinions, and suggestions linked directly to the material.”

The OATS note-taking interface was not originally designed as a discussion forum, and some comments noted the interface differences between OATS and other discussion forums, “… iHelp lends itself to a broader array of unrelated topics, whereas OATS tends to result in more focused discussion on a narrower set of topics. One could see this as good or bad I suppose, depending on the circumstance.” The interface limitations led other users to find problems: “It is not a community chat/messaging system like ihelp”, “[there are] not enough people participating in the discussions.”, and “[it is] Difficult to see whole discussion without clicking on all highlights (some of which are empty).” In particular, because notes were originally intended for more of a personal reminder, rather than for discussion, the interface has limitations. In particular, users have no awareness of where discussions are occurring and so must search highlights manually to find comments.

The interface for discussion did allow some users to discover advantages over typical threaded discussion forums like iHelp, which we hadn’t considered: “It's easy to see where the interest has been concentrated which could give way to potential discussion. You don't have to read the whole paper in order to comment on a highlight, it is simple to become involved without spending hours reading hierarchical posts on i-Help.”, and, “… It is also more conducive to group interaction I think because people tend to rate their interest in a particular iHelp thread based on the title somebody gave it. The titles can be a bit misleading sometimes I find.” So, while the current interface has some limitations our participants discovered some interesting advantages.

Based on the input of our participants, it seems, “Perhaps some kind of familiar forum view”, would be a logical next step. We are finding that most use cases we encounter are for instructors who want, as one participant put it, “focused discussion on a narrower set of topics”. We believe this is quite attractive to instructors, as it allows them to treat a reading as a first-class artefact on which to focus discussion.

We see discussion as a necessary part of the future OATS; however, some new functionality would be required. We need to incorporate annotation-based SNS indicators (such as in AnnotatEd) in the system to allow users to have an awareness of where (and perhaps the amount of) discussion is taking place. Further, OATS will need a traditional interface for discussions; this could be achieved by tying into an existing forum system such as iHelp.

4.6.5.2 System Comparison and Usability

Examining each of the prominent system features allowed us to identify interesting new benefits of such a system for focused discussion. Participants felt overall that the system was useful for the case study task; most users felt it was between slightly and quite useful. We also had users compare OATS with iHelp directly, since all participants had extensive experience using the iHelp system and it has a typical interface among discussion forums. The results were less conclusive in direct comparison. Most users felt that it was only slightly likely that OATS would allow them to more quickly and easily complete the reading and discussion tasks. However, most users were convinced that OATS was more useful for the task, and that it was better than iHelp for reviewing ideas. We feel these impressions would be vastly improved by implementing the features as described in the previous section.

With regards to the usability of OATS in terms of its ease of use, clarity of interactions, and it being easy to learn, we were pleased with the results. We had focused on making OATS as simple as possible using a consistent popup window interface. Participants found the overall usability high; all users responded with a level agreement above the neutral level for these aspects. We will carry our interaction techniques forward with subsequent development on the system.

4.7 **Summary**

Based on our previous work in applying collaborative tagging to e-learning (as described in section 0), we have created OATS, which provides an infrastructure to embed collaborative tagging with other functionalities in online learning materials. The system has served as a proof-of-concept for a unique collection of social software which also illustrates how such functionalities can be decoupled from learning management systems. This architecture allows our system to be easily integrated into any different web-based LMS, relatively easily.

Through in-depth analysis of a multi-week case study in a university class, discussing online reading materials, we have confirmed our development path and intuitions about the usability and usefulness of the system. Further, we have made several interesting discoveries about how interactions and organization should be embedded in e-learning systems. Of particular promise is the highlighting functionality of OATS which has extended previous work on embedding social navigation support into a normally socially disconnected environment. Further, we see research potential around combining social support with discussion forums.

With regards to tagging in our system, our results were somewhat mixed. Overall, tagging was warmly received but severely underused. We attribute this mostly to its lack of usefulness to achieve the goals of the case study. This has provided us with new intuitions about the motivational requirements of tagging users. In particular, we feel we need to provide an environment where tags may persist for some time in order for the overhead of providing tags to be overcome by the recall and organizational benefits of browsing and searching using tags (for further discussion please see section 6.2).

Chapter 5

Collaborative Tagging and Structured Metadata

The creation of adequate metadata remains a significant road-block in much of the current educational technology research endeavours. Such endeavours have lofty goals such as the dynamic assembly of a course using existing learning objects. However, even more modest research goals, such as learning object recommendation, are far from having more than limited success. The problems that the e-learning community face are the same that have arisen around the Semantic Web vision. Both communities continue to define rigid and complex specifications for metadata that will directly enable certain potential functionalities. As we will explain in this chapter, continued research along these lines ignores the substantive evidence that taxonomies and ontologies alone will not achieve the desired outcomes.

While some of our ongoing work at University of Saskatchewan looks at the generation of this metadata as an automatic process through techniques such as data-mining [12][13][60], it still has a way to go before realization. We also regard complementary human generated metadata or annotations as an important resource: directly (human-created metadata used in an un-altered state for some pedagogical purposes) and indirectly (human-created metadata used as a resource in a reasoning process for some pedagogical purposes).

We subscribe to the notion that metadata is best created if it focuses on a particular goal, is contextualized to a particular user, and is created in an ambient manner by observing the actions and interactions of students in learning environments. We believe collaborative tagging and other social annotations have a strong possibility of being a leading method by which we collect this learner-centric metadata. Research continues to emerge that confirms these intuitions [40][63][64][72].

This chapter identifies how the pros and cons of collaborative tagging complement those of structured metadata such as ontologies. We continue by outlining the state-of-the-art in combining the approaches of social software with structured metadata representations in a new field of socio-semantics [65]. Socio-semantics research has the goal of enabling more intelligent applications (through making use of information from social software), which may be able to better achieve purposes such as the accurate recommendation of relevant learning materials.

5.1 Representing Knowledge: Comments on Taxonomies, Ontologies, and Collaborative Tagging

The key to the Semantic Web is the creation of machine consumable knowledge [9]. The main approach being taken is to use ontologies to describe concepts, the properties of concepts, and the relationships between concepts. Among the e-learning community much focus has been placed on investigating how ontologies might facilitate the creation of adaptive e-learning systems. The approach has been to use instances of concept descriptions to annotate resources (often learning objects), as seen in [25][57][66]. However, ontologies are not without their drawbacks. Materialized, these drawbacks not only limit the amount and quality of ontological metadata created but also who can be involved in its creation and therefore the overall usefulness of the approach.

Ontologies extend taxonomies by further describing the relationships between any concepts, effectively turning a hierarchy into a directed graph. Gruber defined an ontology, as "... an explicit specification of a conceptualization" [38], which was refined by Borst as "...a formal specification of a shared conceptualization" [11]. Though ontologies give the foundations for interchangeable knowledge representation, we point out that this does not necessarily entail their conceptualizations being shared. This is one of the most common misconceptions of Semantic Web critics, who make the realization that conceptualizations will rarely share common ground in practice. This has been recognized early on by the Semantic Web community; however, they may have overestimated their ability to define common grounds in mappings between different ontological conceptualizations.

A further issue is that ontology terms can be easily misinterpreted for many reasons, such as synonymous concept definitions, misinterpretation of concepts or different applications in end use. While it is true that a representation of conceptualization may be shared, it is incorrect in saying that the conceptualization itself is "shared". Rather, for a conceptualization to be truly shared, we must also fully represent how it is to be used. Collaborative taggers tend to create metadata with their own end uses in mind. So, collaborative tagging by its very nature lends itself to allow for sharing conceptualizations for achieving shared purposes.

5.1.1 Problems with Current Learning Object Metadata

The IEEE's Learning Object Metadata (LOM) [48] is the most widely used specification for learning object metadata. We examine the LOM here to illustrate the problems with rigid definitions. To fit into our discussion, we will define the LOM as: *a taxonomy of terms, some of which are constrained by preset vocabularies*. The LOM is a big and well designed specification, and has in the neighborhood of 80 terms, but it is at the same time too broad and too specific to be applied to all possible annotations. As described by Brooks et. al. in [13], the LOM requires a lot of information, but at the same time there may be information not adequately represented. There is significant effort used to create and define a large structure of detailed information (represented in the LOM), and the training of instructors to provide good metadata when learning objects are being created has been an ongoing issue. Even the creators of the standard, the IMS Global Learning Consortium, note that the desire does not exist in practice to support such a detailed structure:

> Many vendors expressed little or no interest in developing products that were required to support a set of meta-data with over 80 elements [and the] burden to support 80+ meta-data elements on the first iteration of a product is too great for most vendors to choose to bear. [49]

Further, we have seen in less detailed metadata that non-expert annotators desire help from experts while creating their metadata and this is especially true of relational types of data that are inherent in ontologies [35]. We can extend our discussion of taxonomic metadata, therefore, to the case of ontologies. The difficulties associated with metadata can also be problems for ontologies. Compounding the difficulty of using ontological metadata is the requirement that the annotator has an in-depth knowledge of the use of tools and technologies. Thus, there exists a need to have ontological metadata more easily created and easily used, while making help available throughout the authoring process.

5.1.2 Enter Collaborative Tagging

Although not nearly as expressive for machine knowledge as ontologies, collaborative tags are more easily applied to resources because of their simplicity. Their lack of expressiveness can be seen in more complex domains - for instance, some del.icio.us users have implemented a hierarchy of tags (effectively taxonomies) by pre-pending categorical symbols to tags [39]. This ad hoc extension to collaborative tags still allows ambiguous semantics (e.g. is the child tag a more narrow term than the parent, or are elements within the child a subset of the parent?) and begins to become less useful for other users who are not using a similar strategy. Further, collaborative tags have no property or predicate relationship explicitly implied by the tags. This is okay for humans who are good at determining which keyword implies which relation, but a computer would have great difficulty doing the same thing. For instance, tagging the URL for “Adaptive Hypermedia 2006 Conference” webpage http://ah2006.org with “semantic”, “web”, “Barry” and “Smyth”. Humans seeing these tags can determine fairly easily what the tags mean, especially if they know anything about AH2006. They imply that the conference has some involvement with Semantic Web technologies, and that Barry Smyth is involved in its organization. We can see that computers may have a great deal of trouble determining what a human can see relatively easily. Even so, the second relation may be a bit of stretch even for a human, because of other possibilities for the implied predicate (e.g. “presenter at”). Further, the computer may not be able to group the tags together effectively, “semantic” and “web”, should really be treated as the compound term, “semantic web”. These issues limit the approach greatly, and not just for computers they create problems for human understading too.

Without a hierarchy, tags associated with resources can grow to unruly proportions, even for the tags of a single user. However, the lack of relational information and the ease of simply using keywords without design allows users to annotate more and worry less - if they miss an appropriate tag it is likely that someone else in the community will get it. The universal applicability of this approach arises despite what the majority in a tagging community would deem erroneous tags or meta-noise. Rather, it is this normally useless information which allows the metadata to cater to minorities as well (as described in detail in Section 1.2.2.1). Shirky, on his well read posting, argues for the organic classification system of tags, over rigidly dictated ontological techniques, "... del.icio.us has no idea what the tags mean. The tag overlap is in the system, but the tag semantics are in the users. This is not a way to inject linguistic meaning into the machine." [74] His point hits directly at the main problem with collaborative tagging: the semantics of the tags are only useful for human consumption and computational entities can provide only limited reasoning in a collaborative tagging system. This is in some ways a catch-22: collaborative tagging harnesses the power of the community and has been shown effective in creating large amounts of metadata quickly, but this metadata is limited to human consumption only, which can be quickly overwhelmed by the breadth of tags. Ontologies (in the semantic web sense), on the other hand, were designed specifically to provide rich machine decidable semantic representations. These semantics are found within the subject-property-object relations that have been defined by the W3C's (World Wide Web Consortium) core ontological Resource Description Framework (RDF) technologies. This technique of knowledge representation allows for more intelligent searching and reasoning over resources, as can be seen in [33].

The idea of combining community knowledge sharing techniques with writing RDF statements has resulted in the award winning Confoto[29]. Confoto, like flickr creates annotations of pictures; however, it uses RDF statements instead of keywords. Unfortunately, the application falls short of reaching the ease of use of flickr. Further, there is no way to extend the knowledge represented in the application, as its annotations are based on a closed set of ontological terms, and there is no authoring support for metadata creators.

This examination of ontologies and collaborative tags closes with a brief examination of the process for creating metadata with each approach. Steps in a standard approach might include:

1. learning languages for writing ontologies (RDF, RDFS, and possibly OWL)
2. designing an ontology for the domain or finding and customizing an existing domain ontology
3. creating metadata with instances that fit the ontology
4. assessing the ontology's usefulness for some application(s)
5. returning to the design step when the ontology doesn't meet all needs

The de facto standard for creating metadata with collaborative tagging is much simpler:

1. write appropriate tags (keywords) possibly taking into account community recommendations (suggested tags), until you think you have enough

---

[29] See http://www.confoto.org

5.2 **Socio-Semantics: A Meeting of the Ways?**

Our research and the recent literature have led us to believe that the implicit and explicit information about social networks that exists in social software provides deeper knowledge than what is visible on the surface. These ideas have been shared by certain researchers around the world as they look for practical ways to extract or create knowledge representations to be used by the next generation of computer applications.

The notion of socio-semantics has emerged as an area of research based on interest in leveraging knowledge from social software systems. We simply define socio-semantics as, *extracting and representing deeper knowledge from software systems where communities of users share knowledge*.

We identify three possible approaches to achieving this goal: 1) social support for manual ontology development, 2) social creation of structured knowledge, and 3) analysis of social sources of information.

1) Social Support of Manual Ontology Development and Refinement

As previously discussed there is a strong requirement for author support in metadata authoring. Techniques such as indirect social navigation support seem in particular a very practical method for evolving knowledge (i.e. such as the evolving knowledge bases existing in Wikipedia or del.icio.us) to be incorporated into more structured metadata. Our current work is the only project we know of looking at achieving the use of social authoring support from social software sources.

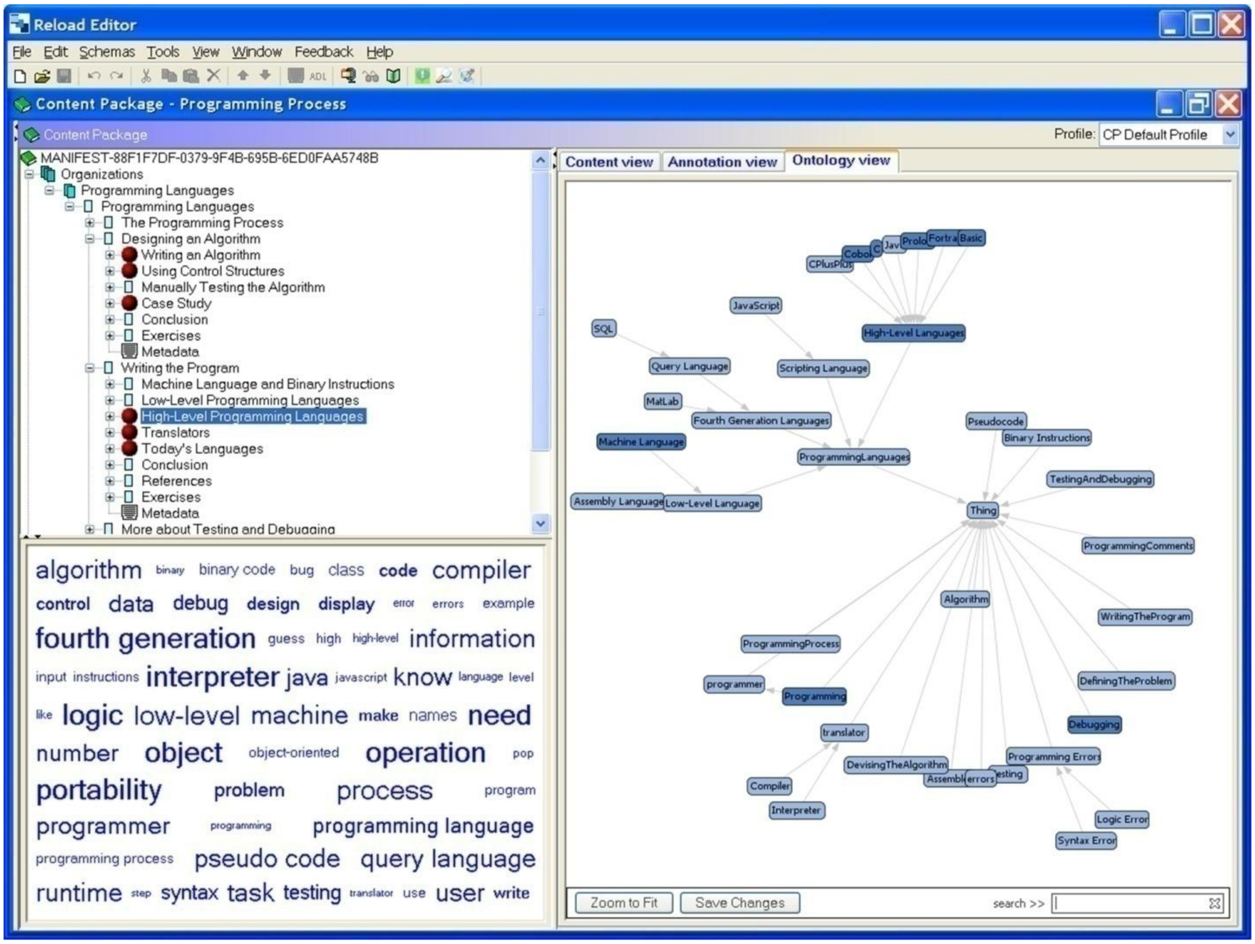


Figure 5.1 A new prototype system to provide social support for ontology enrichment based on community support for tags and measures of semantic relatedness. The lower left pane shows a tag cloud used to advise the ontology author. The right pane, shows the ontology to be extended using the social support from the tag cloud visualization.

We have developed a new prototype system to investigate this new approach (shown in Figure 5.1). The prototype allows tagging data to be displayed to educational metadata authors to enrich ontological representations of their course ontologies. Our system displays a tag cloud visualization of tags applied on related learning content by learners. The size of the tags in the tag cloud indicates the number of times a tag has been applied. We supplement the support information with an additional mapping based on measures of semantic relatedness (a natural language processing technique), using Wikipedia as a corpus. As an ontology author interacts with an ontology, associated tags are displayed in the tag cloud and become darker as semantic relatedness increases, indicating to the author a possible relation. Authors could then select a particular tag in the system as being a new sub-concept or an alternative label for a concept in the domain ontology. We are in the process of evaluating how such support systems based on information from social software systems may facilitate ontology refinement.

2) Social Creation of Structured Knowledge

The social creation of structured knowledge implies altering the knowledge representation process in existing social software systems to create a more structured form of knowledge (i.e. one that could be processed by machines), or by representing knowledge from social software systems in a more structured way.

Our previous work on a framework, called CommonFolks [7], described a method to alter the collaborative tagging process to force users to disambiguate tags based on a taxonomy or ontology of general terms. This process would allow tags to be understood and applied automatically. While this approach was straight-forward, we came to believe the process required too much of the general user, and so we have abandoned this effort.

These intuitions about CommonFolks, were confirmed by the work of Roy Lachica, who developed a website called, fuzzzy.com. Fuzzzy, was different from CommonFolks in that it did not require users to disambiguate tags from the outset; rather users could explicitly relate tags to have semantics when they wanted (for example a user could say the tag “chapter” was a “part of” the tag “book”). The motivation for the user, to engage in this semantic authoring task of relating tags, was to get more specific search results, since tags could have been explicitly disambiguated and related. However, Lachica found, after a small community of users had formed around the website's use, that very few people employed the semantic authoring capabilities (< 1% of the community engaged in this type of authoring) [55].

Similar approaches for authoring semantic data have been more widespread with semantic wikis, such as OntoWiki [3]. These systems extend wikis with the ability to reference and link data, based on an ontology. However, based on our experience we believe these will be short-lived since the effort required of authors is too high.

Other more practical approaches are underway which pull data from social software sources and represent them directly in ontological forms. Most notably dbpedia is a project that mines data from Wikipedia and keeps it in an ontology database. This allows users to ask sophisticated queries of the database, such as "List all 19th century poets from England." This more simple approach seems a better route, since it leverages a huge community of metadata authors and effectively extracts and represents their contributions [4].

3) Analysis of Social Sources of Information

We feel that the most interesting, in terms of potential implications, is automatically extracting data from social software sources using data mining techniques. While this approach is perhaps not the most immediately practical, it takes into account the need for social software systems to remain simple for users. In particular, leveraging information based on social network analysis seems to hold much promise [64] and continues to be under researched. For instance, one could mine del.icio.us for tags, and use statistical data mining approaches of system features, natural language processing of documents and tags, and social network analysis to create taxonomical or ontological representations for sub-communities of del.icio.us users.

There has been some effort in this direction already, though the results have been modest first steps at best. For instance, the most well known tool for semi-automatic ontology generation is TextToOnto [58], which analyses a text corpus and arranges mined keywords into a hierarchical structure representing broader and narrower terms. This work has been refocused by others interested in harvesting folksonomic sources. The conjecture is that since keywords (tags representing important concepts) are already provided from human sources, the machine algorithms may focus on structuring and linking the keywords together based on the more extensive contextual information available in collaborative tagging systems (such as in implicit social networks), to create a more definitive representation than that of analyzing text alone.

We will mention a couple of recent works of note. The work of Van Damme et. al. [79] proposes the use of diverse resources which are first processed and then integrated using existing computational techniques such as basic text processing, statistical analysis, social network analysis, clustering, and ontology mapping. Their work, while seemingly practical and well thought out, only proposes a semi-automatic approach for the creation of a lightweight ontology and still lacks any implementation. This work, though, serves to illustrate a comprehensive approach that incorporates both statistical and knowledge-based reasoning approaches to potentially overcome the limitations of either approach applied alone.

The work of Al-Kalifa and Davis [1] has shown a simple approach which extracts tags from del.icio.us and automatically maps them to a domain ontology. While Al-Kalifa and Davis have an implementation, the simple evaluation describes a study based on searchers using folksonomy-based searching versus searching based on the structured metadata created using their approach. While they show some very modest improvements in the ability to test subjects to locate search targets, the experimental design fails to convince that their particular approach provides any significant advantage.

Given the modest results obtained so far in research in this area, it might seem unreasonable to assume that the automatic creation of structured metadata from social sources holds much hope. However, consider the success and power of the Google search engine. While the Google search algorithms are anchored in latent semantic analysis (a purely statistical technique), they too have used outside information based in the ideas of social networks. The PageRank algorithm which is at the heart of Google's success [35] incorporates ideas such as the weighting of search results using hyperlinks as an indication of the importance of a page. This idea is akin to tags, as the hyperlinks (human created pointers) indicate how important a page is (as do the number of tags would in describing a page on del.icio.us). Tags, though, incorporate more information than does a simple hyperlink, which represents a human created textual description of a web resource. Thus, intuition would suggest that additional information can be harnessed in some way to create better search results. This is, of course, of particular interest to the semantic web community, who are struggling to obtain a real breakthrough based on human-based knowledge descriptions, and thus social software systems, in particular collaborative tagging systems, seem to be a practical direction towards that goal. We believe at the very least that tags will provide useful information for new statistical approaches which take into account human-based voting and knowledge, using algorithms similar to PageRank.

5.3 **Summary**

We have explained how collaborative tagging, which suffers from many problems in terms of its application as metadata, also provides many advantages which may be able to supplement structured metadata. In particular, recent research has set out three main approaches for leveraging socially created data for more structured representations. We feel that given the different approaches, those most likely to succeed will leverage the current level of contributions of social software users. It has been demonstrated that users will not contribute metadata without good reason. Therefore, tasks must remain simple and the payback for users must be immediate, or at least payback must be available in the foreseeable future. This holds true for e-learning, and would be a significant challenge since the e-learning domain usually has strong time constraints (e.g. the length of a course). We feel that overcoming this is possible, by incorporating tagging as required pedagogical activity of students or by allowing access to e-learning course materials to persist beyond the lifetime of a course and even beyond the student's lifetime of the course. While both these solutions are plausible in actual practice there remain major hurdles for the widespread uptake of tagging as a key educational activity. First, for instructors and administrators to be convinced to incorporate tagging as a pedagogical activity we would first need to be shown that tagging is a beneficial activity not just for student's but for the upkeep and use of learning materials. In the case of allowing continuing access to learning materials beyond the lifetime of a course, instructors and administrators would have to agree that the benefit of tagging would outweigh potential risks to intellectual property and copyrights of the materials.

We feel the most interesting approach, that is the automatic analysis of social sources of information, remains largely uninvestigated. While, we know people are linked together by their communities of interest, we are not sure how far this can be extended to modeling the actions and needs of users.

Chapter 6

Conclusion

6.1 Summary of Research and Contributions

Although the individual chapters of this thesis have had several different focuses, they have all centered on one idea, "Collaborative tags represent practical metadata which can supplement other forms of metadata, particularly in e-learning where sufficient metadata is lacking." Collaborative tagging is more than simply keywords, but it includes a process of simple metadata creation and use.

Collaborative tagging has appeared within the last five years, with relatively little fanfare, and yet now can count its users in the millions. So, when we started this research over two years ago, tagging was half as old. Many questions existed that first needed to be investigated about the technology itself before we would know how to apply it in a domain such as e-learning. Through our research we have extended the state of the art about how tagging may be applied in new situations, how tags can be visualized, how we can interact with tags, how we can extend tagging to new interface actions, and how we can represent tagging as structured knowledge.

Chapter 1 described collaborative tagging in detail, by placing it in terms of other methods of classification in practice and in research. We presented seminal research which has made concrete intuitions about the complex dynamics of collaborative tagging.

In Chapter 2 we described the common methods of exploring and interacting with tags. Through, our analysis we were able, with some success, to confirm and discover new properties of a visualization technique that previously had been largely considered too subjective for such empirical study. We have also provided a basis and a study-paradigm for future empirical studies of tag clouds.

In Chapter 3, we contextualized tagging within the domain of e-learning providing questions of interest for research. We followed this by providing a first analysis of tagging by learners, which has provided us with interesting knowledge of how tagging might be best used in e-learning. Further, we have new research directions for overcoming the cold-start issue in new social software communities.

In Chapter 4, we applied our knowledge of the domain of e-learning and research experience in social software to create a unique system for e-learning, the Open Annotation and Tagging System. We described how the architecture of OATS can be replicated to allow such "value-added" systems to be easily integrated into many different e-learning systems. Finally, we reported the results of an exploratory case-study, which allowed us to learn about the motivational requirements of tagging, while providing interesting directions for research in the extension of the tagging metaphor, to include collaborative highlighting. Further, we have learned the benefits and implications of seeding discussion directly in learning material.

In Chapter 5, we outlined the state-of-the-art in a new field of research, socio-semantics, which aims at leveraging socially created knowledge to help create more structured forms of metadata. We have, for the first time, defined the different approaches and outlined how our research, current and past, fit into this promising new field.

6.2 Collaborative Tagging: E-Learning versus the Open Web

Given our contributions to research in understanding collaborative tagging and in particular its use in e-learning the main questions that arise are "What conclusions can be drawn from this thesis with regards to collaborative tagging in e-elarning? Is collaborative tagging appropriate in e-learning and how should it be applied?"

While we have provided some intuitions that e-learners can add keywords and create typical patterns of consensus and learners may see some value to tagging in their learning activities, we have fallen short of definitively showing when and how collaborative tagging in particular may be successful in e-learning. Tagging systems on the open Web (those that are freely accessible and persistently available to any web surfer) in particular provide value. The value is that the tags will be around for the foreseeable future, and therefore have applicability and value for the foreseeable future, and this provides users with intrinsic motivation to tag. This was not the case in our study of the OATS system, where tagging was not persistent (limited until the end of the course). Further, the task was not sufficiently complicated that learners need to use tags to categorize information; rather they found an ability to achieve the study goals and their own goals mainly through browsing highlights. This is not the case for users on the Open Web, where they are better motivated to tag a resource since it is a quick and easy way back to find a single resource. Therefore, the cost of adding a tag is outweighed by the benefit of being able to easily rediscover the resource for the foreseeable future.

Our OATS case study suggests that tagging must have a higher benefit-cost ratio than those presented in short-term and simple scenarios. While tagging is easy and of extremely low cost (in terms of effort), the fact remains people will not continuously tag for free. We have cited two possible ways to overcome this issue: by adding persistent access over time, not only to tags but access to learning resources; and/or by having tagging be a pedagogical activity. Persistent access to resources on the Open Web is usually expected and thus can be used in e-learning, but it may be a more contentious issue for instructors and administrators to continuously share their personal learning materials (which they may never have intended to be widely and continuously available). This would be based on individual desires and goals of instructors and students on a case-by-case basis. While tagging would certainly add considerable benefit if widely available, say to a student throughout an undergraduate degree, it is difficult to imagine a situation where tagging would be widely applicable across all courses and instructors would widely accept the value of incorporating tags in their learning materials. In regards of tagging as a pedagogical activity there remain several issues which would need to be investigated: the effect tagging has on knowledge gain, how to incorporate tagging into and draw from learning theories, and how tagging can be incorporated into day-to-day instruction. Further do instructors see the added benefits of pedagogical reflection, learner reflection, metadata collection, and learner self-organization as outweighing the costs of the added learner effort, pedagogical effort (in terms of integrating the technology in a course), and the potential loss of control provided by social navigation support (the risk remains that the system would offer poor advice).

While, we continue to believe tagging to have particular benefits to e-learners, it would seem that the most likely situation, given all the potential hurdles of either incorporating the technology as a pedagogical activity or making tagging continuously available, is to simply study it in cases where it is directly applicable. For instance, an applicable situation would be a course which involves sharing and discussion of online articles.

Can collaborative tagging supplement metadata in e-learning? We continue to believe the answer is “yes”, but very careful consideration must be paid to apply tagging in the correct situation. In particular there must be a particular problem which tagging will solve (most likely organizing and sharing online resources), and the problem must be sufficiently complex that the learners need a solution. Such situations would allow the motivation for tagging to be intrinsic in its use by students, and therefore the implications could be studied further in an authentic situation.

6.3 **Future Work**

The Ecological Approach [60] proposes that metadata may be discovered and used by analyzing the fine-grained interactions of learners in their digital environments, as they set out to achieve different learning goals (e.g. recommending a learning object of the appropriate cognitive level and length, given the abilities and time constraints of the learner). We now believe that the social dimension is a key to supplementing rich interaction traces to identify and disambiguate the purposes of learners. While, the ecological approach proposes automatic support for achieving pedagogical goals, I propose an extension where the user interface and the data discovery process are inseparable parts of the approach, and that support in both cases are available through the emergent communities using the system, just as in collaborative tagging systems.

Such a system would observe the actions of the users and based on heuristics classify successful and unsuccessful actions. By associating the actions with those of a discovered community of interest, recommendations could be made to a user on what action or path leads to success, but also other related courses of action. These ideas are inspired by combining the notions of the Ecological Approach with those of social navigation support.

For instance, consider a new graduate student interested in methods of evaluation in human computer interaction. Where would he begin? Assume he is familiar with human computer interaction, but has forgotten the names of all of the methods of evaluation. If he could find a review paper it would be a great reference. Searching for “human computer interaction evaluation review” on Google fails to reveal any appropriate papers immediately, but maybe there is a link to such a paper somewhere in the results (or maybe not). He could spend time clicking link after link, or he could wait until the next day to ask his fellow students. But surely somebody has searched for this before, or something very similar. There may be another student in North America doing this right now. What if we could have captured the student’s actions in his search and associated it with other searches and searchers? What if we could capture enough about the student and his actions to help him in some way?

What collaborative tagging facilitates, which search engines do not, is serendipitous discovery of resources of interest. By pivot browsing on the username of a user who seems to be interesting, a set of new terms not yet considered can reveal a wealth of new information. Can search and social navigation support be combined to give automatic recommendations? How much information can we capture about a user through their actions in order to find an appropriate clique for them within a community? How many of the successes and failures of other users can we meaningfully capture? How can recommendation be best incorporated into a users' everyday search actions?

My future research will continue to be inspired by the social software movement, and continue to look at what can be learned and discovered from emergent communities. Social software has shown that "intelligent systems" can be inherently simple, but that they must rely on understanding how people will use the systems. My work will strive to understand how a little bit of algorithmic intelligence and an understanding of human-computer interactions can lead to more intelligent social software.

References

[1] Al-Khalifa, H. S., Davis, H. C. and Gilbert, L. Creating Structure from Disorder: Using Folksonomies to create Semantic Metadata. In Proceedings of the *3rd International Conference on Web Information Systems and Technologies* (WEBIST). March, 2007.

[2] Anderson, C. The Long Tail. *Wired*. October, 2004.

[3] Auer, S., Dietzold, S., and Riechert, T. A Tool for Social, Semantic Collaboration. In Proceedings of *5th International Semantic Web Conference*. November 5th - 9th, Athens, GA, USA. pp. 736–749. 2006.

[4] Auer, S., Bizer, C., Lehmann, J., Kobilarov, G., Cyganiak, R., and Ives, Z. DBpedia: A Nucleus for a Web of Open Data. In Proceedings of *6th International Semantic Web Conference*, November 11th –15th, Busan, Korea. 14 pages. 2007. (to appear)

[5] Bateman, S., Farzan, R., Brusilovsky, P., and McCalla, G. OATS: the open annotation and taggind system. In Proceedings of the *3rd Annual E-learning Conference on Intelligent Interactive Learning Object Repositories*, November 9th – 11th, Montreal, Canada. Available at http://www.lornet.org/Default.aspx?tabid=644. 10 pages. 2006.

[6] Bateman, S., Brooks, C., McCalla, G., and Brusilovsky, P. Applying Collaborative Tagging to E-Learning. In the Proceedings of the *Workshop on Tagging and Metadata for Social Information Organization*, held in conjuction with the *16th International World Wide Web Conference*. May 7, Banff, Canada. Available at http://www2007.org/workshops/paper_56.pdf. 7 pages. 2007.

[7] Bateman, S., Brooks, C., and McCalla, G. Collaborative Tagging Approaches for Ontological Metadata in Adaptive E-Learning Systems, in Proceedings of *Workshop on Applications of Semantic Web Technologies for e-Learning*, at the *4th International Conference on Adaptive Hypermedia and Adaptive Web-based Systems*. June 20th – 23rd, Dublin, Ireland.pp. 3-12. 2006.

[8] Bauer, A., and Koedinger, K.R. Selection-Based Note-Taking Applications. In the proceedings of the *ACM Conference on Human Factors in Computing Systems 2007 (CHI '07)*. pp. 981 – 990. 2007.

[9] Berners-Lee, T., Hendler, J., and Lassila, O. The Semantic Web. *Scientific American* Vol. 20, pp. 34–43. 2001.

[10] Bloom, B.S. *Taxonomy of educational objectives, handbook I: the cognitive domain*. David McKay Co Inc., New York. 1956.

[11] Borst, W.N. *Construction of Engineering Ontologies*. PhD thesis, University of Twente, Enschede, NL. Centre for Telematica and Information Technology. 1997.

[12] Brooks, C., and McCalla, G. Towards Flexible Learning Object Metadata. *International Journal of Continuing Engineering and Lifelong Learning*. Vol. 16, No. 1/2, pp. 50-63. 2006.

[13] Brooks, C., McCalla, G., and Winter, M. Flexible learning object metadata. In Proceedings of *3rd International Workshop on Applications of Semantic Web Technologies for E-Learning*, Amsterdam, The Netherlands. Held in conjunction with the *12th International Conference on Artificial Intelligence in Education*. 8 pages. 2007.

[14] Brooks, C. H., and Montanez, N. Improved Annotation of the Blogosphere via Autotagging and Hierarchical Clustering. Proceedings of the *15th International World Wide Web Conference*. Edinburgh, Scotland. pp. 625-631. 2006.

[15] Brusilovsky, P. KnowledgeTree: a distributed architecture for adaptive e-learning. In *Proceedings of the 13th international World Wide Web Conference on Alternate Track Papers & Posters* May 19th – 21st. New York, NY, USA.). ACM, New York, NY, 104-113. 2004.

[16] Brynjolfsson, E., Hu, Y. and Simester, D. Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales. February 2007. Online article vailable at http://ssrn.com/abstract=953587. Accessed November 12th, 2007.

[17] Bumgardner, J. *Building Tag Clouds in Perl and PHP*. O'Reilly Media 2006.

[18] Cadiz, J. J., Gupta, A., and Grudin, J. Using Web annotations for asynchronous collaboration around documents. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work* Philadelphia, Pennsylvania, United States.. ACM, New York, NY. pp. 309-318. 2000.

[19] Cardinaels, K., Meire, M. and Duval, E. Automating metadata generation: the simple indexing interface. in Proceedings of *14th International World Wide Web Conference*. Chiba, Japan. pp. 548-556. 2005.

[20] Chiarella, A., and Lajoie, S. Enabling the collective to assist the individual: CoREAD, a self-organizing reading environment. In Proceedings of *World Conference on E-Learning (E-Learn 2006)* October, Honolulu, HI. AACE, 2753-2757. 2006.

[21] Chisholm, W., Vanderheiden, G., and Jacobs, I. Web Content Accessibility Guidelines 1.0 W3C Recommendation 5-May-1999. W3C Recommendation available at http://www.w3.org/TR /1999/WAI-WEBCONTENT-19990505

[22] Coupland, D. *Microserfs*. HarperCollins, Toronto, 1995.

[23] Crane, D., Pacarello, E., and James, D. *Ajax in Action*. Greenwich, CT: Manning, 2006.

[24] Dieberger, A., and Lönnqvist, P., Visualizing Interaction History on a Collaborative Web Server, in Proceedings of *Eleventh ACM Conference on Hypertext and Hypermedia*, pp. 220-221. 2000.

[25] Dolog, P., Henze, N., Nejdl, W., and Sintek, M. The personal reader: Personalizing and enriching learning resources using semantic web technologies.. In Bra, P.D., Nejdl, W., (eds). Proceedings of the *Third International Conference Adaptive Hypermedia and Adaptive Web-Based Systems*, August 23-26, 2004, Eindhoven, The Netherlands. Volume 3137 of Lecture Notes in Computer Science. Springer. pp. 85–94. 2004.

[26] Dourish, P., and Chalmers, M. Running Out of Space: Models of information navigation, in Proceedings of the *ACM Conference on Human Factors in Computing Systems (CHI 1994)*. 10 pages. 1994.

[27] Dubinko, M., Kumar, R., Magnani, J., Novak, J., Raghavan, P., and Tomkins, A. Visualizing tags over time. In Proceedings of the *15th International World Wide Web Conference*. pp. 193-202. 2006.

[28] Farzan, R., and Brusilovsky, P., AnnotatEd: A Social Navigation and Annotation Service for Web-based Educational Resources. In Proceedings of Proceedings of *World Conference on E-Learning (E-Learn 2006)* pp. 2794-2802. 2006.

[29] Farzan, R., and Brusilovsky, P. Social Navigation Support through Annotation-Based Group Modeling. In Proceedings of the 10th International Conference on User Modeling, pp. 463-472. 2005.

[30] Fielding, R.T. *Architectural Styles and the Design of Network-based Software Architectures*, Doctoral dissertation, University of California, Irvine. 2000.

[31] Friedman, T. L. It's a Flat World, After All. *New York Times Magazine*. April 3, 2005.

[32] Friesen, N. What are educational objects? *Intereactive Learning Environments.* Vol 9, No. 3. 2001.

[33] Gasevic, D., Hatala, M.: Searching web resources using ontology mappings. In proceedings of the *Integrating Ontologies Workshop* held in conjunction with the 3rd *International Conference on Knowledge Capture*. pp. 33-40.22005.

[34] Golder, S., Huberman, B.A.: Usage patterns of collaborative tagging systems. *Journal of Information Science.* Vol. 32. pp.198–208. 2006.

[35] Google. Google Technology. Webpage available at http://www.google.com/technology/ accessed on December 7th, 2007.

[36] Greenberg, J., and Robertson, W.D. Semantic web construction: An inquiry of authors' views on collaborative metadata generation. In Proceedings of the *International Conference on Dublin Core Metadata for e-Communities*, Florence, Italy, Firenze University Press. pp. 45–52. 2002.

[37] Grossman, Lev. Time's Person of theYear: You. *Time Magazine*. Time Inc. Dec. 13, 2006.

[38] Gruber, T.R. A translation approach to portable ontology specifications. *Knowledge Acquisition*. Vol. 5. pp.199–220. 1995.

[39] Guy, M., and Tonkin, E. Tidying up tags? *D-Lib Magazine.* Vol. 12. 2006.

[40] Halpin, H., Robu, V., and Shepherd, H. The Complex Dynamics of Collaborative Tagging. In the Proceedings of the *16th International World Wide Web Conference*. Banff, Canada. pp. 211-220. 2007.

[41] Halvey, M., and Keane, M. T. An Assessment of Tag Presentation Techniques. In the proceedings of the *16th International World Wide Web Conference*. Banff, Canada. pp. 1313-1314. 2007.

[42] Hammouda, K. and Kamel, M. Automatic Metadata Extractor. Project page available at http://pami.uwaterloo.ca/projects/lornet/software/ame.php, accessed January 30th, 2007.

[43] Hammouda, K. and Kamel, M. Data mining in e-learning, Chapter in *Proceedings E-Learning Networked Environments and Architectures: A Knowledge Processing Perspective*, Springer. 2005.

[44] Hassan-Montero, Y., and Herrero-Solana, V. Improving tag-clouds as visual information retrieval interfaces. In the Proceedings of *International Conference on Multidisciplimary Information and Technology (InfoSciT2006)*. 6 pages. 2006.

[45] Hill, W. C., Hollan, J. D., Wroblewski, D., and McCandless, T. Edit wear and read wear. In Proceedings of the *ACM Conference on Human Factors in Computing Systems* Monterey, California, United States, May 03 - 07. ACM Press, New York, NY. pp. 3-9. 1992.

[46] Hill, W.C., and Hollan, J.D. History-enriched digital objects. *In Third ACM Conferenceon Computers, Freedom and Privacy*. San Francisco, CA. pp. 917-920. 1993.

[47] Hendler, J. The Dark Side of the Semantic Web. *Intelligent Systems*, Vol. 22, No. 1. pp. 2-4. 2007.

[48] IEEE, Inc. *Draft standard for learning object metadata*. IEEE P1484.12.1-2002, (2002).

[49] IMS Global Learning Consortium Inc. *IMS meta-data best practice guide for IEEE 1484.12.1-2002 standard for learning object metadata, version 1.3*. Available at http://www.imsglobal.org/metadata/mdv1p3pd/imsmd_bestv1p3pd.html. 2004.

[50] John, A. and Seligmann, D. Collaborative tagging and expertise in the enterprise. In the Proceedings of *Collaborative Web Tagging Workshop* held in Conjunction with *15th International World Wide Web Conference*. Edinburgh, Scotland. 6 pages. 2006.

[51] Kasser, O., & Lemire, D. Tag-Cloud Drawing: Algorithms for Cloud Visualization. Proceedings of the *Tagging and Metadata for Social Information Organization Workshop* in conjunction with *16th International World Wide Web Conference (WWW '07).* 10 pages. 2007.

[52] Koivunen, M.R. Annotea and semantic web supported collaboration. Invited talk at *Workshop on User Aspects of the Semantic Web* in conjunction with the *European Semantic Web Conference*. May, Heraklion, Greece. 13 pages. 2005.

[53] Kucera, H., and Francis, W. N. *Computational Analysis of Present Day American English*. Brown University Press, Providence. 1967.

[54] Kurhila, J., Miettinen, M., Nokelainen, P., and Tirri, H. EDUCO - A collaborative learning environment based on social navigation. In the Proceedings of *Second International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*, pp. 242-252, 2002.

[55] Lachica, R. Towards holistic knowledge creations and interchange Part 1: Socio-semantic collaborative tagging. Talk presented at the *Third International Conference on Topic Maps Research and Applications*. October 11-12. Leipzig, Germany. Slides available at http://www.informatik.uni-leipzig.de/~tmra/2007/slides/lachica_TMRA2007.pdf. 2007.

[56] Larson, K. The Science of Word Recognition or how I learned to stop worrying and love the bouma. Online article accessed Sept. 8, 2007. Avaialable at http://www.microsoft.com/typography /ctfonts/WordRecognition.aspx. 2007.

[57] Lytras, M.D., Pouloudi, A., and Korfiatis, N. An ontological oriented approach on e-learning: Integrating semantics for adaptive e-learning systems. In Proceedings of the *11th European Conference on Information Systems (ECIS 2003)*. Naples, Italy 16-21 June. Avaialable at http://is2.lse.ac.uk/asp/aspecis/20030090.pdf. 11 pages. 2003.

[58] Maedche, A., and Staab, S. Semi-automatic Engineering of Ontologies from Text. In the Proceedings of the *12th International Conference on Software Engineering and Knowledge Engineering*. 2000.

[59] Mathes, A. Folksonomies - cooperative classification and communication through shared metadata. Online article available at http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html. Accessed September 13, 2007. 2004.

[60] McCalla, G. The Ecological Approach to the Design of E-Learning Environments: Purpose-based Capture and Use of the Information about Learners *Journal of Interactive Media in Education (JIME)*. Vol. 1. Special Issue on the Educational Semantic Web. pp. 18. 2004.

[61] Millen, D., Feinberg, J., and Kerr, B. Dogear: Social Bookmarking in the enterprise. in Proceedings of *ACM Conference on Human Factors in Computing Systems (CHI 2006)*. Montreal, Canada. pp. 11-120. 2006.

[62] Millen, D., and Feinberg, J. Using Social Tagging to Improve Social Navigation. In the Proceedings of the *Workshop on Social Navigation and Community-Based Adaptation Technologies* In Conjunction with *Adaptive Hypermedia and Adaptive Web-Based Systems*. June 20th, Dublin, Ireland. 10 pages. 2006.

[63] Mika, P. *Social Networks and the Semantic Web*. Springer, 2007.

[64] Mika, P. Ontologies are us: A unified model of social networks and semantics. In Proceedings of the *4th International Semantic Web Conference,* volume 3729 of Lecture Notes in Computer Science, pages 522–536. Springer, 2005.

[65] Morville, P. *Ambient Findability*. O'Reilly Media, 2005. ISBN 978-0-59-600765-2.

[66] Nejdl, W., Wolf, B., Staab, S., and Tane, J. Edutella: Searching and annotating resources within an rdf-based p2p network. In Frank, M., Noy, N.F., Staab, S. (eds) the Proceedings of the *International Workshop on the Semantic Web*, Hawaii, May 7, 2002. Volume 55 of CEUR Workshop Proceedings., CEUR-WS.org. 6 pages. 2002.

[67] O'Reilly, T. What Is Web 2.0 - Design Patterns and Business Models for the Next Generation of Software. Online article available at http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html. Accessed Sept. 13, 2007. 2005.

[68] Palmatier, R.A., and Bennet, J.M. Note-taking habits of college students. *Journal of Reading.* Vol. 18, pp. 215-218. 1974.

[69] Papert, S. Constructionism: A New Opportunity for Elementary Science Education. NSF Grant Application. Accessed November 13th, 2007. Available at http://nsf.gov/awardsearch/showAward.do?AwardNumber=8751190. 1987.

[70] Postman, N. Five Things We Need to Know About Technological Change. Transcript of speech given at *NewTech Conference*, Denver, CO, USA. March, 1998. Available at http://www.mgi.polymtl.ca/marc.bourdeau/InfAgeTeaching/OtherLinks/Postman.pdf accessed November 13, 2007.

[71] Rivadeneira, A.W., Gruen, D.M., Muller, M.J., and Millen, D.R. Getting our head in the clouds: toward evaluation studies of tagclouds. In the proceedings of the *ACM Conference on Human Factors in Computing Systems (CHI 2007)*.pp. 995-998. 2007.

[72] Sen, S. , Lam, S., Rashid, A., Cosley, D., Frankowski, D., Osterhouse, J., Harper, M., and Riedl, J. tagging, community, vocabulary, evolution. In the Proceedings of *Conference on Computer Supported Cooperative Work (CSCW '06).* November, 2006. Banff, Canada. pp. 181-190. 2006.

[73] Shared Web Annotation & Content Aware Services, Gibeo, Available at http://gibeo.net, 2006.

[74] Shirky, C. Ontology is overrated: Categories, links, and tags. http://www.shirky.com/writings/ontology_overrated.html. Accessed January 30th, 2007. 2005.

[75] Sinclair, J., and Cardew-Hall, M. The folksonomy tag cloud: when is it useful? *Journal of Information Science*. Cardew-Hall , May. 2007.

[76] Suignar, M. CSS Text Level 3 W3C Working Draft 6 March 2007. W3C Working Draft available at http://www.w3.org/TR/2007/WD-css3-text-20070306/. 2007.

[77] Svensson, M., Laaksolahti, J., Höök, K., and Waern, A. A recipe based on-line food store. In the Proceedings of the *5th International Conference on Intelligent User Interfaces* New Orleans, Louisiana, United States. ACM Press, New York, NY, 260-263. 2000.

[78] Tufte, E., R. *The Visual Display of Quantitative Information.* Graphics Press, Cheshire, Conn. 2001.

[79] Van Damme, C., Hepp, M., and Siorpaes, K. FolksOntology: An Integrated Approach for Turning Folksonomies into Ontologies. In the Proceedings of the *Workshop on Bridging the Gap Between Semantic Web and Web 2.0* collocated with *4th European Semantic Web Workshop*. 2007.

[80] Van der Wal, T. "Folksonomy". Online article available at http://www.vanderwal.net/folksonomy.html Accessed Sept. 8, 2007. 2007.

[81] Ware, C. *Information Visualization: perception for design*. Academic Press, London, 2000.

[82] Wexelblat, A., and Mayes, P. Footprints: History-rich tools for information foraging. In Proceedings of the *ACM Conference on Human-Computer Interaction*. pp. 270-277. 1999.

[83] Wiley, D. A. Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. Online article available at http://reusability.org/read/chapters/wiley.doc Accessed November 13th, 2007. 2000.

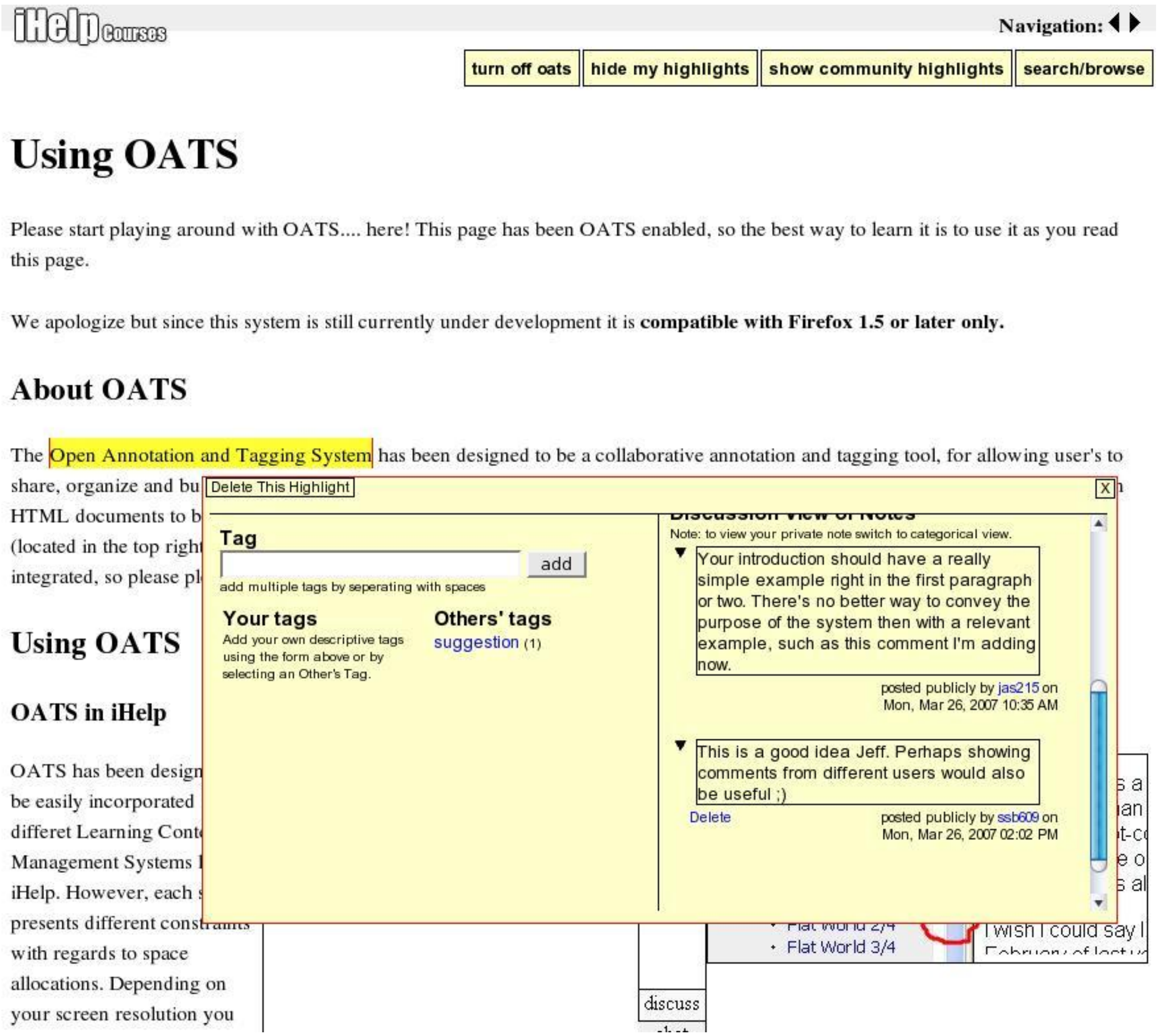
[84] Williams, R.L., and Eggert, A.C. Note-taking in college classes: Student patterns and instructional strategies. *The Journal of General Education*. Vol. 51, No.3, pp. 173-199. 2002.

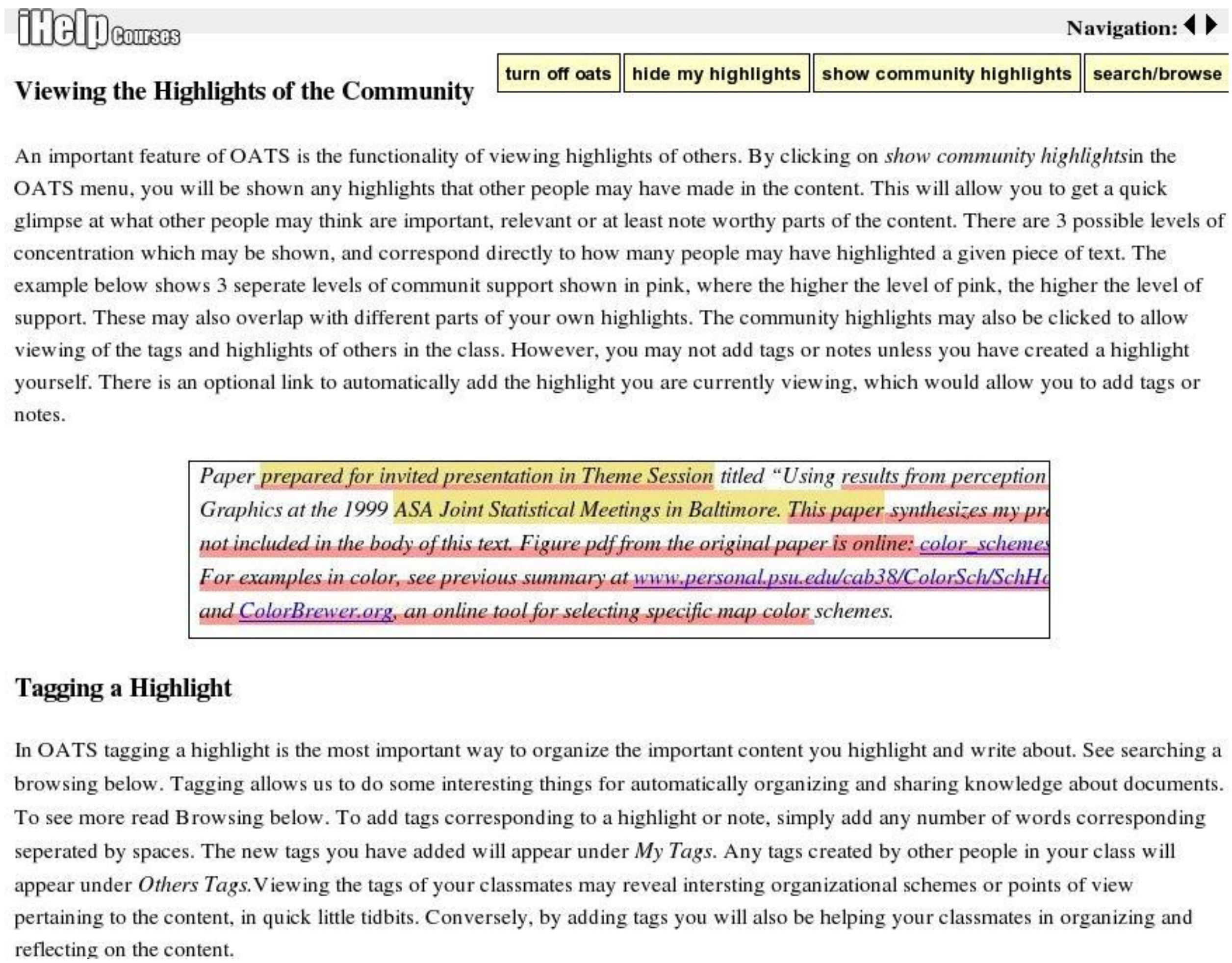
[85] Xin, C., and Glass, G. Enhancing online discussion through web annotation. in G. Richards (Ed.), In the Proceedings of *World Conference on E-Learning (E-Learn 2005)* October, Vancouver, Canada. pp. 3212-3217. 2005.

[86] Zeldman, J. Tag clouds are the new mullets. Online article available at http://www.zeldman.com/daily/0405d.shtml. Accessed Sept. 8, 2007. 2005.

Appendix A

Introductory Webpage for the OATS Study





iHelp Courses Navigation:

turn off oats | hide my highlights | show community highlights | search/browse

Viewing the Highlights of the Community

An important feature of OATS is the functionality of viewing highlights of others. By clicking on *show community highlights*in the OATS menu, you will be shown any highlights that other people may have made in the content. This will allow you to get a quick glimpse at what other people may think are important, relevant or at least note worthy parts of the content. There are 3 possible levels of concentration which may be shown, and correspond directly to how many people may have highlighted a given piece of text. The example below shows 3 seperate levels of communit support shown in pink, where the higher the level of pink, the higher the level of support. These may also overlap with different parts of your own highlights. The community highlights may also be clicked to allow viewing of the tags and highlights of others in the class. However, you may not add tags or notes unless you have created a highlight yourself. There is an optional link to automatically add the highlight you are currently viewing, which would allow you to add tags or notes.

> *Paper prepared for invited presentation in Theme Session titled "Using results from perception Graphics at the 1999 ASA Joint Statistical Meetings in Baltimore. This paper synthesizes my pr not included in the body of this text. Figure pdf from the original paper is online: color_schemes For examples in color, see previous summary at www.personal.psu.edu/cab38/ColorSch/SchH and ColorBrewer.org, an online tool for selecting specific map color schemes.*

Tagging a Highlight

In OATS tagging a highlight is the most important way to organize the important content you highlight and write about. See searching a browsing below. Tagging allows us to do some interesting things for automatically organizing and sharing knowledge about documents. To see more read Browsing below. To add tags corresponding to a highlight or note, simply add any number of words corresponding seperated by spaces. The new tags you have added will appear under *My Tags*. Any tags created by other people in your class will appear under *Others Tags.*Viewing the tags of your classmates may reveal intersting organizational schemes or points of view pertaining to the content, in quick little tidbits. Conversely, by adding tags you will also be helping your classmates in organizing and reflecting on the content.

Appendix B

OATS Exit Survey

OATS Survey

Thank you for participating in the development of OATS. As a final request and for the qualitative and quantitative assessment of OATS we ask you to complete this short questionnaire. This questionnaire should take you no more than 10 minutes to complete, and will be invaluable to the continuing work on the OATS system. While, answering all questions would be very useful to us, please complete as much as convenient for you. Most questions are just checkboxes, and so won't take much time.

Of course, the conditions of the consent form still apply. All information will be annonymized and in no way be able to identify you.

Any questions or concerns can be sent to Scott Bateman <scott.bateman@usask.ca> or Prof. Gord McCalla <mccalla@cs.usask.ca>

User Demographics

This information will give us an idea about you, your background, and your experience.

1. NSID [ ]
2. What year of study have you completed? [ ]
3. In managing and studying reading materials (e.g. books, notes, print-outs of slides, etc.) during your University career do you use or have you ever used: (check all that apply)
   - ☐ highlighting in the reading material
   - ☐ marginalia (writing in the margins) in the reading material

☐ writing keywords or concept words; in the reading material itself or using a separate medium

☐ writing summaries; in the reading material itself or using a separate medium

☐ some **other** technique for helping you study, organize or learn (please describe below)

If **other** please describe:

4. Have you ever used a highlighting annotation system other than OATS, such as in a word processor, Adobe Acrobat or another web based system?

   ○ Yes ○ No

If Yes, which ones and how frequently did you use them?

If No, were you aware of such functionality being available in other programs?

○ Yes ○ No

5. Before using OATS, did you use *tagging with keywords* for personal organization?

   ○ Yes ○ No

   a. **If Yes**, which websites, how often and for what purposes? Please briefly describe how you use *keyword tags*.

If No, were you aware of the use of *keyword tagging* for personal organization?

○ Yes ○ No

Other than using OATS and iHelp, have you ever used a *discussion forum or message board*?

○ Yes ○ No

b. **If Yes**, which systems/websites, how often, and for what purposes? Please briefly describe how you use *discussion forums and/or message boards*.

If No, were you aware of any web-based discussion forums and/or message boards, which are used for support on a wide range of topics?

○ Yes ○ No

Usefulness of OATS

This information will give us an idea of how and what you found easy when using OATS to accomplish a specific **task**.

The **task** in the 408 course was to interactively discuss important issues. This was divided into two parts, the first used iHelp Discussion, and the second used OATS. In the following please compare the **task** using OATS and iHelp.

1. In general I would find OATS useful to accomplish the task.

likely	○	○	○	○	○	○	○	**unlikely**
	extremely	quite	slightly	neither	slightly	quite	extremely	

2. Using OATS would allow me to *accomplish the task more quickly* than in iHelp.

likely ○ ○ ○ ○ ○ ○ ○ **unlikely**

extremely quite slightly neither slightly quite extremely

3. Using OATS would make it *easier for me to accomplish the task* than in iHelp.

likely ○ ○ ○ ○ ○ ○ ○ **unlikely**

extremely quite slightly neither slightly quite extremely

4. Using OATS would allow me to *more easily review my ideas and the ideas of others* for the task than in iHelp.

likely ○ ○ ○ ○ ○ ○ ○ **unlikely**

extremely quite slightly neither slightly quite extremely

5. I would find OATS *more useful* than iHelp to accomplish the task.

likely ○ ○ ○ ○ ○ ○ ○ **unlikely**

extremely quite slightly neither slightly quite extremely

3. Feature of OATS

Please rate following features of OATS with regard to their usefulness for you.

1. Being able to highlight text (the yellow highlights you created in the text) was:

useful ○ ○ ○ ○ ○ ○ ○ **unuseful**

extremely quite slightly neither slightly quite extremely

2. Being able to view the highlights created by others (the pink highlighted text which you could turn on through "View Community Highlights") was:

useful	○	○	○	○	○	○	○	**unuseful**
	extremely	quite	slightly	neither	slightly	quite	extremely	

3. Being able to add discussion notes based on the highlights was:

useful	○	○	○	○	○	○	○	**unuseful**
	extremely	quite	slightly	neither	slightly	quite	extremely	

4. Being able to "tag" (add keywords) to the discussion notes and highlights was:

useful	○	○	○	○	○	○	○	**unuseful**
	extremely	quite	slightly	neither	slightly	quite	extremely	

5. What makes OATS useful?

6. What makes OATS not useful?

7. Can you think of any other purposes other than the task set in class that OATS would be useful or well-suited for?

Usability of OATS

Please indicate with what likelihood you could make each of the following statements.

1. Learning to use OATS was easy for me.

   likely ○ ○ ○ ○ ○ ○ ○ **unlikely**

   extremely quite slightly neither slightly quite extremely

2. My interaction with OATS was clear and understandable.

   likely ○ ○ ○ ○ ○ ○ ○ **unlikely**

   extremely quite slightly neither slightly quite extremely

3. I found OATS easy to use.

   likely ○ ○ ○ ○ ○ ○ ○ **unlikely**

   extremely quite slightly neither slightly quite extremely

4. What makes OATS easy to use?

5. What makes OATS not easy to use?

6. Can you think of anything that would improve how easy OATS is to use?

Final Comments

Do you have any final comments or thoughts on OATS that were not previously mentioned in this survey or in email communications?
(i.e. desired features, particular problems, related ideas, etc)

Submit Completed Survey | Reset Survey