Facette: Using Facets to Improve Tag-based Bookmarking

by

Peter Lai

Submitted to the Department of Electrical Engineering and Computer Science

In partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 2009

[September 2009]

Author .....

Department of Electrical Engineering and Computer Science

August 26, 2009

Certified by .....

David R. Karger

Professor of Computer Science and Engineering

Thesis Supervisor

Accepted by .....

Dr. Christopher J. Terman

Chairman, Department Committee on Graduate Students

Facette: Using Facets to Improve Tag-based Bookmarking

by

Peter J. Lai

Submitted to the Department of Electrical Engineering and Computer Science on Aug. 26, 2009,
in partial fulfillment of the requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

Facette is a web service that uses facets to enhance the organizational capabilities of tag-based bookmarking systems. As with other bookmarking services, Facette allows users to associate tags with bookmarks to assist the retrieval of information. Facette also allows users to classify tags through use of facets. To create these facets, Facette introduces a method of facet creation called free faceted tagging. To assist the classification of tags using facets, Facette extends the concept of collaborative tagging and introduces the concept of collaborative faceting.

For ease of implementation, Facette is implemented as an augmentation of Delicious' tagging system. Faceted information is saved to Delicious through use of new tag syntax. Faceted information is retrieved from Delicious through use of Delicious's API.

Since its public debut on February 16th, 2009, Facette has successfully assisted the use of faceted tagging. Facette has been used by several hundred Delicious users. Facette's tagging interface has led to the bookmarking of thousands of pages with thousands of faceted tags. Overall, Facette has successfully encouraged the creation of faceted information.

Thesis Supervisor: David R. Karger
Title: Professor of Computer Science and Engineering

Acknowledgements

I would like to thank my advisor, David Karger, for allowing me to choose my own thesis topic. Two years ago, I first began implementation of Facette as a UROP project with the Haystack group at CSAIL. After completion of an initial prototype of Facette, David Karger accepted my requests to continue development and analysis of Facette for my graduate project. David Karger also provided me with research guidance that enabled the completion of this thesis. The first prototype of Facette contained an amalgamation of various research ideas. David Karger's guidance helped me refine from my initial ideas a core concept that could be studied through analysis of tagging.

I would like to thank the Haystack group at MIT for nurturing my growth as a computer scientist. During my year as a graduate student, the Haystack group provided me with constant feedback on my thesis and provided me with constant testing of Facette. The Haystack group exposed me to some of the best web development talent at MIT. David Huynh was particularly influential in my learning advanced web development techniques. Work with his project, Exhibit, exposed me to the pleasures of client-side web development.

Most importantly, I would like to thank my parents. I'd like to thank my mother for providing me the determination and patience necessary to excel at MIT. I'd like to thank my father for inspiring me with stories of success.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Most popular online bookmarking services use tags to help users label and retrieve bookmarked web pages. Users assign textual tags to pages they encounter and can then search or browse through those tags in order to re-find bookmarked information. At a certain scale, these tags efficiently assist the retrieval of information; however, after a user has created more than several hundred tags, the tags themselves become difficult to navigate. The following figure contains an example tag navigation panel for a user who has created over several thousand tags.

Figure 1.1: The Tag-navigation Panel for a Heavy Delicious User

After the creation of several thousand tags, users encounter the novel challenge of finding relevant tags from within this long list of items.

Some tags used in popular online bookmarking services also convey ambiguous meanings. For example, in certain contexts, the tag 'google' may indicate that a webpage contains topics relating to Google. In other contexts, the tag 'google' may imply that a webpage is product of Google. When a user browses a personal collection of bookmarks, this ambiguity

may not necessarily result in confusion; however, when a user collaboratively browses an aggregation of everyone's bookmarks as is supported by several popular bookmarking services, such ambiguity may result in difficulty finding relevant information. Enhanced organizational techniques are needed in order to prevent these types of ambiguities.

Facette is a web service that adopts concepts from *faceted classification* in order to alleviate organizational issues associated with tagging. Facette provides an interface for navigating bookmarks using tags that have been organized within *facets*. Facette also provides an interface for creating *faceted tags*. To assist the creation of faceted tags, Facette generalizes the concept of collaborative tagging and uses "wisdom of crowds" to suggest which facets someone might want to use and which tags someone might want to place within those facets. Facette has been deployed and has been used by several hundred users to create several thousand bookmarks. Based upon data collected through use of Facette, it appears that some people find value in faceted tagging and that collaborative faceting is effective for choosing facets and values within facets.

As with most online bookmarking services, Facette allows users to associate tags with bookmarks. Facette also allows users to classify tags using organizational groups called facets. The following figure contains a screenshot of Facette's browsing interface.

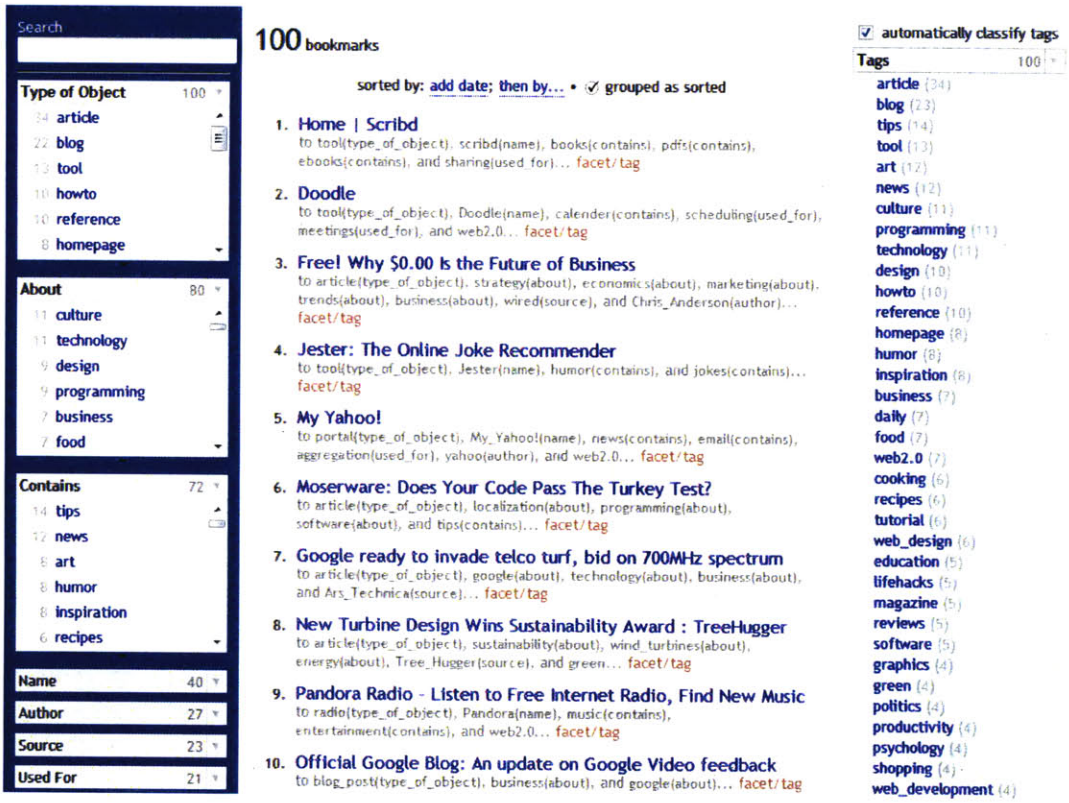

Figure 1.2: Facette's Browsing Interface

In the right-hand column of Facette's browsing interface, a list of tags is shown. This list of tags resembles the lists of tags shown in traditional tag-based interfaces. In the left-hand column of Facette's browsing interface, groups of tags are arranged within facets. These facets provide users with an additional level of organization than enhances the process of guided navigation. Faceted navigation is a well studied technique in information retrieval [1]. Concepts from faceted navigation have been incorporated within several popular commercial websites [2].

The contributions of this thesis focus upon generating faceted information. To create faceted information, Facette introduces a method of facet creation called free faceted tagging. When bookmarking a page, a user is given the ability to create an arbitrary set of facets. The following screenshot shows the tagging interface that Facette uses for bookmarking a page.

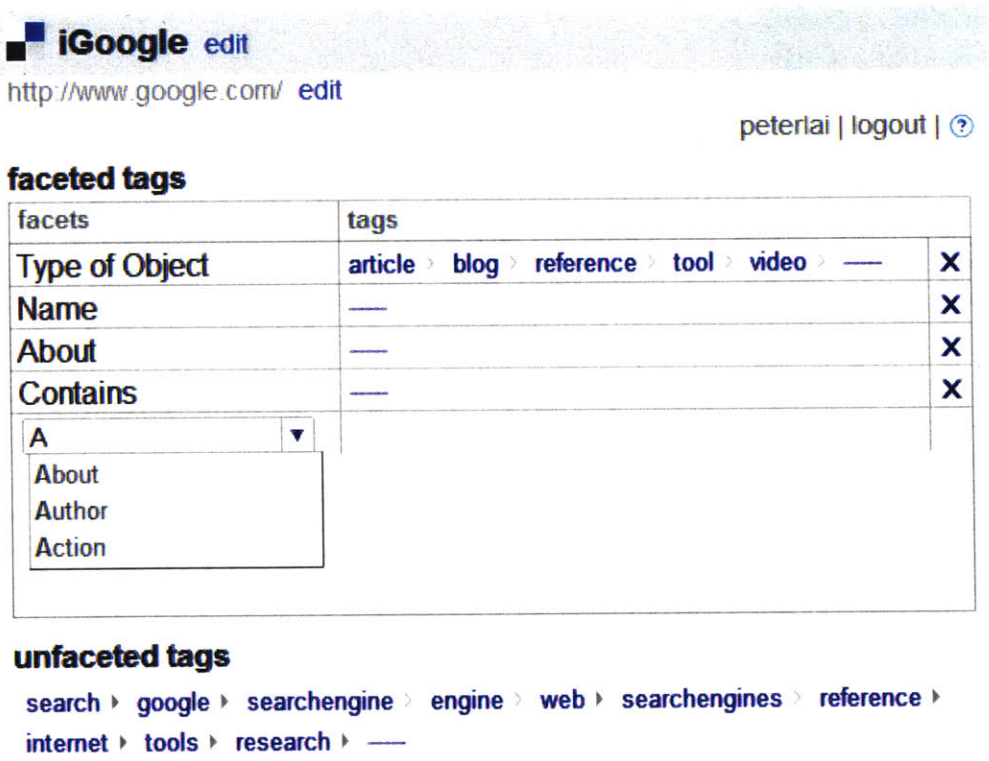

Figure 1.3: Facette's Tagging Interface

In this screenshot, a facet creation textbox has been activated. A facet is created for each phrase submitted using this textbox. The dropdown menu shows a list of values for facets; however, the user is free to enter arbitrary text.

Free faceted tagging also allows the user to abstain from using facets. Certain tags do not easily fit within facets. For example, many users have difficultly creating a facet for the tag 'web2.0'. To encourage the generation of faceting information without discourage the creation of tags, Facette allows users to create unfaceted tags. The section label 'unfaceted tags' in Figure 1.3: Facette's Tagging Interface shows the part of Facette's tagging interface that allows users to created unfaceted tags. This part of Facette's interface is modeled after the interface of existing tagging services and should seem familiar to many users.

Finally, Facette introduces the concept of collaborative faceting. Collaborative tagging is a process in which suggested tags are generated based upon previously created tags [3].

Collaborative faceting is a natural extension of this concept. With collaborative faceting, suggestions for facets and tags are generated through analysis of previously-created faceted tags. Collaborative faceting generates several types of suggestions. Collaborative faceting is capable of generating suggestions for facets and tags. The screenshot below shows one of the ways in which Facette assists faceted tagging.

Figure 1.4: Facet Suggestion Dropdown for Tags

In this screenshot, a dropdown for the tag 'google' suggests a list of facets that this tag may belong to. Because the tag 'google' has often been placed within the 'Name' facet by other users, the 'Name' facet is shown as one of three suggested facet for this particular tag.

Free faceted tagging in combination with collaborative faceting enables the creation of faceted information within tag-based bookmarking systems. The benefits of these methods will be discussed at the end of this chapter, but first, this introduction reviews the weaknesses of tagging, presents the benefits of faceted classification, and evaluates existing methods of generating faceted information using tags.

1.1 Problems with Tagging

Although tags have been praised for successfully encouraging the creation of content-oriented metadata [4], tags have limitations that prevent tag-based systems from solving some common problems in information retrieval. Tags have a limited ability to support guided navigation, and, in certain situations, tagging yields ambiguous metadata.

1.1.1 Limited Navigation

Guided navigation describes the practice of providing navigational cues within a browsing interface to assist the retrieval of information. One form of guided navigation involves the use of tags. When finding bookmarks related to a specific topic, users may use tags to filter lists of items. However, at a certain scale, tags insufficiently support the retrieval of information. To help users navigate a list of tags, a type of interface called a tag cloud is often provided. When a tag within a tag cloud is clicked, a user's collection of bookmarks is filtered for items containing that tag. The following figure contains an example tag cloud for a user with several thousand bookmarks.

Figure 1.5: The Tag-navigation Panel for a Heavy Delicious User

After several thousand tags have been created, a user encounters the novel challenge of finding relevant tags from within this interface. At this scale, a second level of organization for these tags seems necessary. Several professors have noted the necessity of enhancing the organization of tags. Hassan-Montero and Hearst both have studied the deficiencies of tag clouds [5] [6]. Tags within tag clouds are difficult to scan. Without use of additional structure, tags are not well suited for use in scalable interfaces that implement guided navigation.

1.1.2 Ambiguous Semantics

The tags used in tag-based systems lack a semantic precision that is necessary for the maintenance of a controlled vocabulary. Often times, tags are used for different pages to convey different meanings. For example, in certain contexts, the tag 'google' may imply that that a page is a product of Google. In other contexts, the tag 'google' may be used to indicate that a page is about Google. These types of tag ambiguities have been blamed by some researchers for creating chaos within systems that use tagged information [7]. Mathes uses the tag 'filtering' to demonstrate similar tag ambiguity [7]. Some web pages are about filtering and contain topics such as Vodka filtering. Other pages such as Last.FM are used for filtering music and information. Without the presence of additional context, the implied meaning of a tag that has been used to label a page may be difficult to determine.

1.2 Advantages of Faceted Classification

Faceted classification is a well-studied method of improving the organization of information. Faceted classification excels at organizing information that can naturally be described using orthogonal sets of categories [8]. Faceted classification relies upon two primary units of organization: facets and facet values. Facets typically comprise of keywords that describe object properties. Example facets describing a paper include date, creator, topic, and discipline [3]. A list of natural property values could be assigned to each of these example facets. The 'Date' facet likely contains a range of dates that describe when the paper was published. The 'Topic' facet likely contains keywords such as 'economy' or 'psychology'. The use of faceted classification is commonplace on the web [2]. www.newegg.com, www.kayak.com, and www.epicurious.com are example sites that integrate use of faceted classification.

One advantage of faceted classification is its support of faceted navigation. Faceted navigation improves a user's ability to explorer unfamiliar collections of items. When faced with an unfamiliar collection of items, many users on the internet shift into a pattern of browsing called orienteering [9]. Orienteering describes the behavior of using navigational cues to iteratively explore the structure of new information. Users exploring new information often lack

specific targets and often lack effective keywords for search [9]. In lieu of effective keywords for search, users faced with new information depend upon navigational options.

Faceted navigation is a form of navigation that excels at providing users with tools for browsing unfamiliar information [10]. By allowing users to filter information upon multiple dimensions, facet-based browsers provide users with an enhanced sense of control. The presence of multiple movement options reduces the difficulty of exploration. During the orienteering process, users depend upon navigational cues. Each value within a facet offers the user a new browsing option. Finally, when they are faced with too many browsing options, faceted navigation allows users to focus only upon dimensions that are immediately relevant [1].

In review, faceted classification is well suited for organizing multi-dimensional information. Faceted classification provides support for faceted navigation and improves a user's ability to discover information within unfamiliar collections of items.

1.3 Faceted Classification of Tags

When applied to the organization of tags, faceted classification alleviates many of the limitations associated with tagging [11] [12]. The faceted classification of tags improves the specificity of tag metadata. Without information describing how tags relate to items, tags suffer from ambiguity of meaning [7]. The faceted classification of tags helps alleviate this limitation by providing context during interpretation of each tag. The tag 'google' can either apply to the topic of a webpage or the source of a webpage. After this tag has been added to the 'source' facet, this tag's meaning is no longer ambiguous.

The faceted classification of tags also allows tag metadata to better support guided navigation. After being faceted, augmented tags have dimensional information that may be used within standard faceted browsing interfaces. Augmented tags include facets such as 'topic', 'object type', and 'name'. Such facets describe properties and are typical of facets that have been naturally generated through faceted classification [11]. After facets have been generated, augmented tag metadata can be browsed using standard facet-based browsers.

Facette's browsing interface demonstrates several advantages of using faceted tags. The following figure contains a screenshot of Facette's browsing interface.

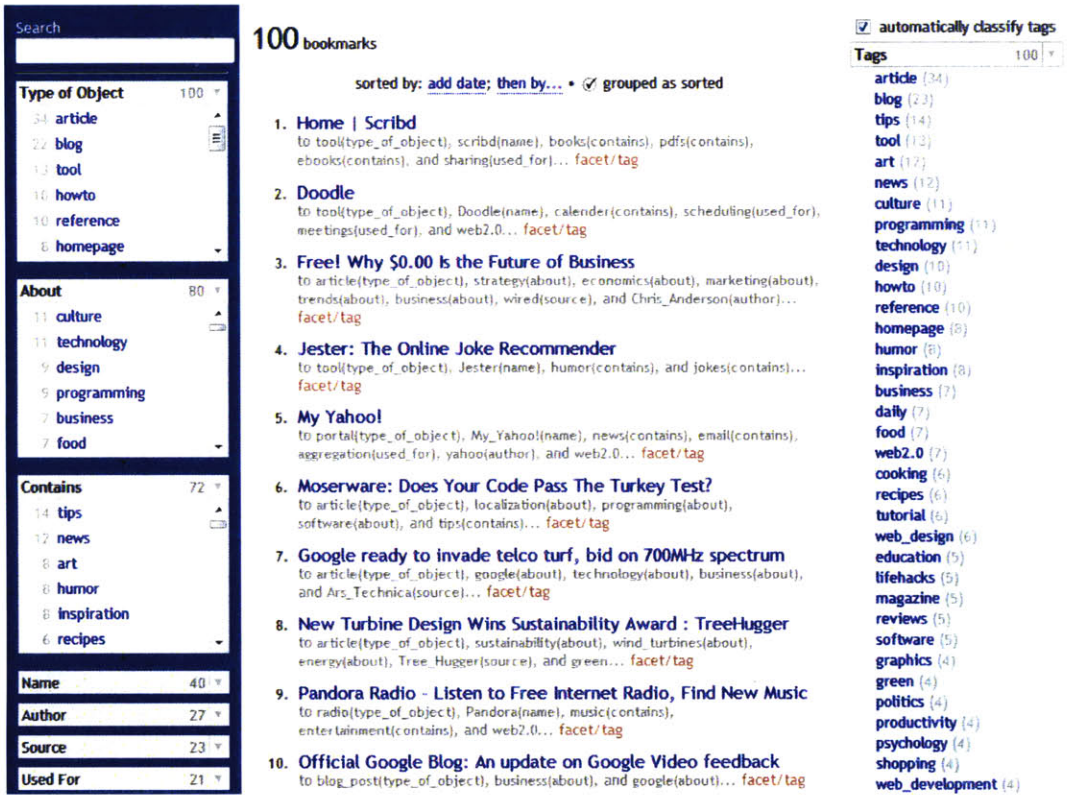

Figure 1.6: Facette's Browsing Interface

In the left-hand column of Facette's browsing interface, facets are used to group together related tags. Within the 'Type of Object' facet, tags that are present include 'article', 'blog', 'tool', and 'howto'. Within the 'About' facet, tags that are present include 'culture', 'technology', 'design', and 'program'. By grouping related tags, Facette's interface addresses the difficultly of scanning tags when several hundred tags are present within the system. Facette's interface also improves the usability of tags with multiple meanings. As previously stated, the tag 'google' can either be a topic or a source. Using Facette's interface, a user can filter information based upon specific uses of that tag. In review, Facette's interface addresses several limitations of tag-based metadata by incorporating use of faceted information.

1.4 Generating Faceted Information

Once faceted information has been generated, the data can easily be loaded into Facette's browsing interface; unfortunately, the generation of faceted information is difficult. There exist several methods for generating faceted information using tags. Methods for generating faceted information are typically either automated or manual. Automated methods for faceting information use statistical techniques such as K-means clustering, subsumption, lexical co-occurrence computation, and inference using WordNet [13]. Manual methods for faceting rely upon the users' creation of tags within facets [12].

1.4.1 Automated Methods

Tags can be used to generate faceted information through statistical methods of automated assignment [14] [15]. Schimitz [14] uses statistical methods of inferring faceted ontologies in order to adopt the benefits of faceted classification while using tag metadata from Flickr [14]. Results from Schimitz's paper indicate that statistical methods of inferred assignment can potentially generate usable metadata. The faceted information generated by their algorithm contains a low percentage of misclassified tags.

A second system uses a related method of tag assignment to automate the grouping of tags [15]. The algorithm in their paper generated results of sufficient quality to be tested experimentally within a popular bookmarking service called RawSugar.

Finally, Stoica and Hearst use a combination of synonym clustering and relationship labeling to semi-automate the process of generating faceted information [13]. Although their algorithm requires human assistance, the authors state that only small adjustments are necessary for the production of acceptable metadata.

1.4.2 Manual Methods

Other methods of generating faceted information involve manual tag classification. Tags can be classified by forcing users to create tags within a predefined set of facets [12] [16]. In a study by

researchers from Bar-lan University, users were required to create tags within one of ten predefined facets [16]. For their experiment, the author states that such constraints generally led to a more comprehensive description of tagged items. A similar method was used in the creation of Facetag [12]. Facetag is a social tagging system that forces users to create tags within a predefined set of facets. The facets chosen for Facetag were based upon recommendations from the Classification Research Group. Results from their user studies indicate that their system successfully adds context to tags and allows faceted navigation.

1.4.3 Problems

The aforementioned methods of faceting allow for significant enhancement of tag metadata; however, none of the aforementioned methods are ideal for generating faceted information using tags. Each method has weaknesses that can be alleviated through iterative improvements.

Methods of automating the generation of faceted information do not address ambiguities introduced during the tagging process. As mentioned before, tags may have multiple meanings. Without input from the user during the tagging process, it becomes difficult to infer the correct meanings of certain tags. For example, if the tag 'google' was used to label a page, the tag ambiguously refers to either the source of the page or the topic of the page. An automated method of faceting would have difficulty resolve this ambiguity. Furthermore, faceted information generated through use of automated methods contains errors that require human correction. Research in automated tag classification primarily focuses upon improving the accuracy of automated assignment [13]. The solution offered by Stoica and Hearst produces acceptable results; however, these results are only obtained after use of human input [13].

Methods of manual faceting have the advantages of improving both the tagging process and the browsing process. Unfortunately, restricted manual faceting hinders the flexibility of tagging and increases the difficult of tag creation. Some tags cannot easily be classified using facets. For example, users have difficulty classifying the tag 'web2.0'. In the study from Bar-Ilan, subjects encountered similar difficulties when forced to place tags within facets [16]. The author notes that some tags could not easily be placed within a single facet and would probably

best be left as a keyword for free text search. In contrast with Bar-Ilan's system, Facette's tagging interface allows tags to remain unfaceted.

Existing methods of manual faceting also fail to provide users with a comprehensive set of facets. By restricting users to a fixed set of facets, the study from Bar-Ilan University prevented users from using ideal facets for the classification of certain tags [16]. The author notes that having a fixed set of facets risks forcing users to fit information within inappropriate facets. The lack of optimal facet options may also discourage users from creating certain tags that cannot be classified using the options that have been provided. To prevent users from encountering such restrictions, Facette's tagging interfaces enables the dynamic creation of facets.

1.5 Contributions

1.5.1 Free Faceted Tagging

In this thesis, I introduce a method of generating faceted information called free faceted tagging that improves upon concepts from existing methods of generating faceted information. Free faceted tagging is a manual method of faceting that allows users to create both tags and facets during the process of tagging. During the creation of tags, users have the option of placing tags within existing facets. If a desired facet is unavailable, the process of free faceted tagging allows for the dynamic creation of the desired facet. In addition, if the faceting of a tag is not possible or is unnecessary, the process of free faceted tagging does not require that the tag be placed within a facet.

Free faceted tagging helps resolve tag ambiguities introduced during the tagging process. Certain tags may have multiple meanings. As mentioned before, the tag 'google' can either describe the source of a page or the topic of a page. By encouraging the use of facets, free faceted tagging encourages the user to be more explicit with how a tag is used. If the tag is within the 'Source' facet, the tag has been used to describe the source of a page. If the tag is within the 'Topic' facet, the tag has been used to describe the topic of a page.

In contrast with existing methods of manual faceting, free faceted tagging allows the user to dynamically create facets. When faceting tags, if a desired facet does not exist, the user may create the desired facet and add tags to this option. For example, when faceting the tag 'weekly', a user may want to use a facet called 'Update Frequency'. By default, this facet does not exist; however, using Facette's facet creation textbox, this facet can easily be added to Facette's tagging interface.

To preserve the flexibility of traditional tagging, free faceted tagging allows the user to abstain from using facets. This flexibility allows the user to use tags that are difficult to facet such as 'web2.0' without forcing the use of inappropriate classification. This aspect of free faceted tagging was introduced specifically to alleviate weaknesses encountered in the study by Bar-Ilan [16].

1.5.2 Collaborative Faceting

In this thesis, I also introduce the concept of collaborative faceting. Collaborative tagging describes the process of generating tag suggestions based upon previously generated tags for a particular page. Collaborative faceting is a natural extension of this idea.

Collaborative tagging is used in several tag-based bookmarking services and provides several benefits. Collaborative tagging has been shown to help with vocabulary convergence. A study by Golder and Huberman cites that, when collaborative tagging has been implemented, users typically select tags that have been suggested by previous users. The authors note that the vocabulary used to describe a page empirically converges within 100 bookmarks [3].

Collaborative tagging also helps simplify the process of tag creation. Collaborative tagging provides users with high quality tags that are likely to be selected for labeling a page. As previously mentioned, when collaborative tagging has been implemented, users typically select tags that have been suggested by previous users [3]. The presence of suggested tags minimizes the effort needed to conjure relevant keywords. For popular pages, a set of relevant keywords already exists. Also, with the implementation of tag selection through use of the mouse,

collaborative filtering reduces the number of keystrokes needed to add tags to a page. Suggested tags can be selected through use of mouse clicks instead of through use of typing.

Collaborative faceting is a natural extension of collaborative tagging. Collaborative faceting is the process of using previously created faceted information to generate suggestions for future users. Collaborative faceting can be used to generate facet suggestions for tags that have not yet been place within facets. Through aggregation of previously created faceted information, collaborative faceting can be used to suggest that the tag 'google' be placed within either the 'Name' facet, 'About' facet, or 'Author' facet. Because several people had previously placed the 'google' tag within these three facets, these three facets have become facet suggestions for this tag. The following figure shows a screenshot of Facette's facet suggestion interface for tags.

Figure 1.7: Facet Suggestion Dropdown for Tags

In this interface, when the menu item for a facet is selected, the tag becomes faceted and is copied into the selected facet.

Collaborative faceting is also capable for generating suggestions for facets that a user may want to create using Facette's facet creation interface. Finally, collaborative faceting is capable of producing suggested faceted tags for pages that have previously been bookmarked using Facette. The details of these two features will be discussed in Chapter 2 of this thesis.

In review, Facette introduces several contributions. Facette introduces free faceted tagging. Free faceted tagging allows the user to create an arbitrary number of new facets and allows the user to abstain from using facets. Facette also introduces collaborative faceting. Collaborative faceting is capable generating facet suggestions and tag suggestions that simplify

the process of faceted tagging. Together, free faceted tagging and collaborative faceting provide several improvements to the process of generating faceted information.

Chapter 2

Facette

2.1 Overview

Thus far, no existing system has implemented free faceted tagging or collaborative faceting. In this chapter, a proof-of-concept system is presented. Facette is a bookmarking service that incorporates concepts from faceted navigation to assist the organization of bookmarks. To browse bookmarked information, Facette uses a browsing interface that displays both tags and facets. Facette's browsing interface preserves the use of tag navigation while simultaneously adding support for faceted navigation.

To create faceted information, Facette uses a custom tagging interface that implements free faceted tagging and uses collaborative faceting. Facette's tagging interface allows users to place tags within facets as tags are created. If no facets are suitable for a tag, the classification of that tag is not required. Facette's tagging interface also enables the creation of new facets if existing facets are not satisfactory.

To assist the process of tagging, collaborative faceting is used to generate faceted tag suggestions and facet suggestions. Facet suggestions are listed for unfaceted tags that have previously been faceted. Facet suggestions are shown when the user attempts to create new facets. Faceted tag suggestions are shown for pages that have previously been bookmarked.

Facette's suggestion algorithm relies upon the processing of user-generated faceted information. Facets from submitted faceted tags are aggregated to create facet suggestions for dynamically created facets and unfaceted tags. The faceted tags for bookmarked pages are aggregated to create faceted tag suggestions for previously bookmarked pages.

For ease of implementation, Facette is designed to be an augmentation of a popular bookmarking service called Delicious. Facette's user data is stored on Delicious. Faceted metadata is stored using an augmentation of Delicious tag syntax. The browsing of collected items is achieved through use of Delicious' API and through use of an existing faceted browsing

interface called Exhibit [2]. A caching system is built into Facette's browser plugin to reduce Facette's dependencies on data transfers.

2.2 Browsing Interface

2.2.1 Tag Navigation

Facette's browsing interface resembles the browsing interface of traditional tag-based bookmarking services. Tag-based bookmarking services use tag navigation to help users browse bookmarked information. To find bookmarks that have been labeled using a specific tag, the user can filter bookmarks by clicking tags within the service's browsing interface. The following figure contains a screenshot of the bookmarking browsing interface for Delicious, a popular tag-based bookmarking service.

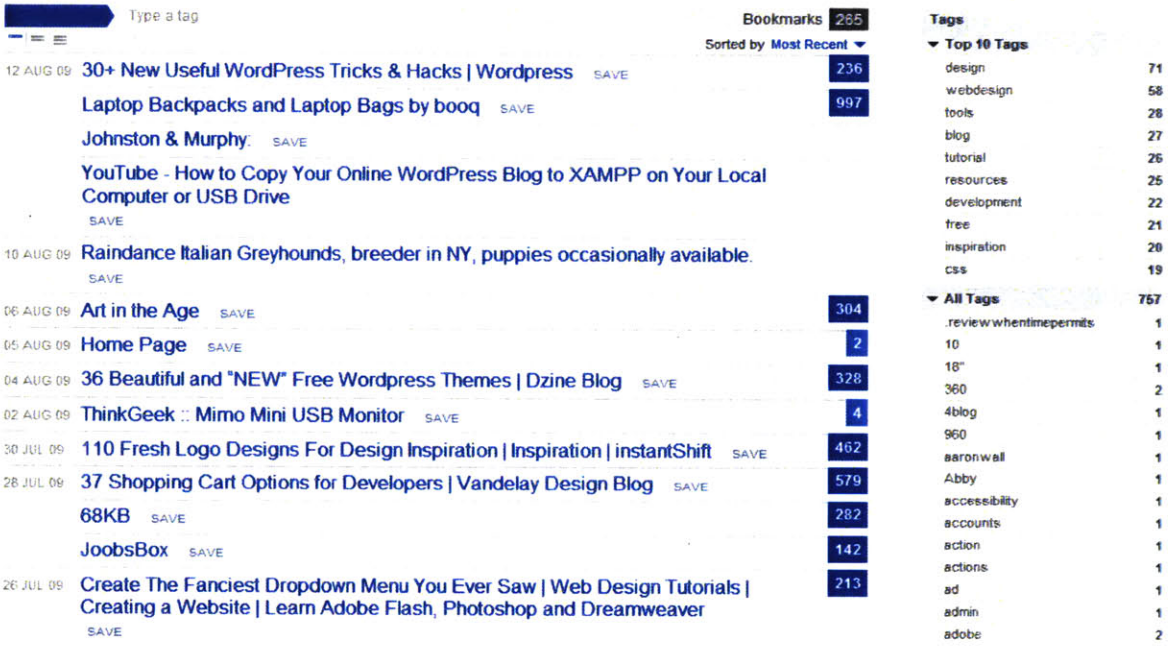

Figure 2.1: Bookmarking Browsing Interface for Delicious

In Delicious's interface, all tags are listed in the right hand column of the page. The clicking of a tag within this column will filter which bookmarks are shown.

Facette's browsing interface preserves support of tag navigation. For consistency with Delicious and other popular tag-based bookmarking services, a list of tags devoid of faceted information is displayed within the right-hand column of the interface. A screenshot of Facette's browsing interface is shown below.

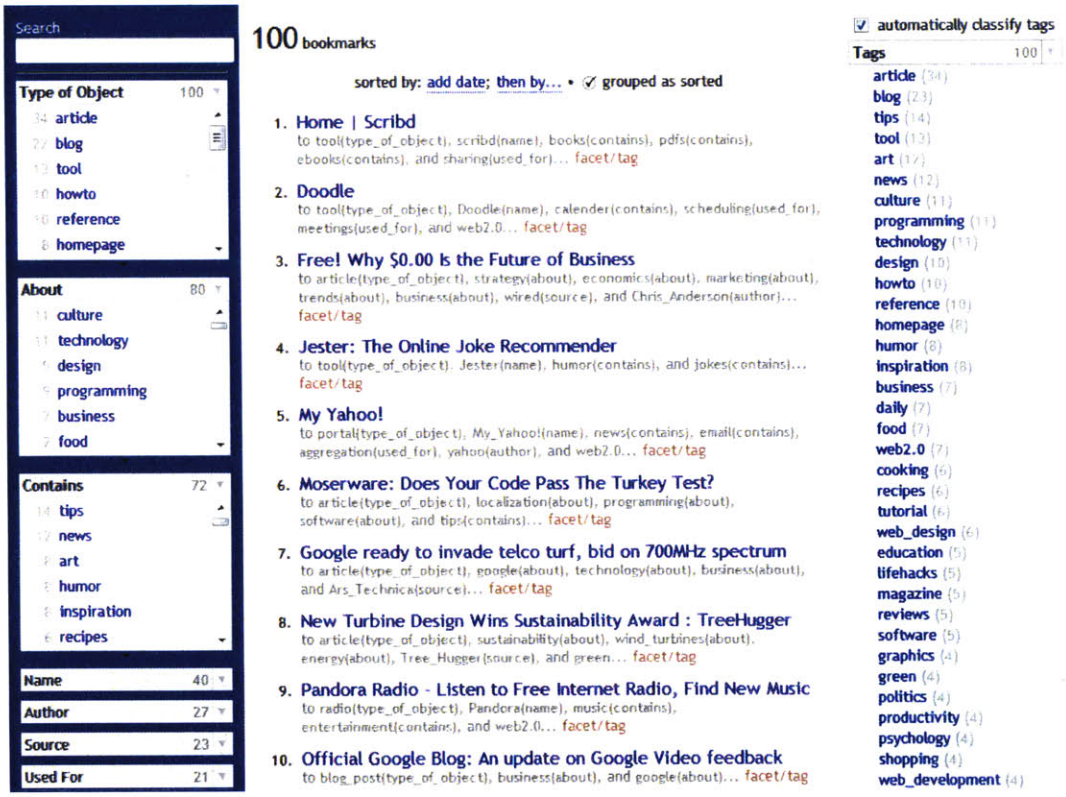

Figure 2.2: Facette's Browsing Interface

This right-hand column of Facette's interface contains unfaceted tags and also faceted tags that have been stripped of faceted information. By including tags with stripped faceted information, this column better supports tag navigation as is implemented in standard tag-based systems. If desired, the user may also select an option to exclude the present of faceted tags within this column.

2.2.2 Faceted Navigation

Facette's browsing interface adds support for faceted navigation. Facet-based controls are shown in the left-hand column of Facette's browsing interface. In Figure 2.2: Facette's Browsing Interface, several facet-based controls are shown. The first control allows for the filtering of bookmarks using values within the 'type of object' facet. The second control allows for the filtering of bookmarks using values within the 'about' facet. As tags within these controls are clicked, the bookmarks in the interface are dynamically filtered. The behavior of Facette's faceted navigation interface is inherited from the faceted browsing framework, Exhibit [2].

2.3 Tagging Interface

Facette's tagging interface is designed to assist the creation of faceted information while maintaining consistency with several aspects of traditional tagging interfaces. To create faceted information, Facette's tagging interface implements free faceted tagging. Facette's tagging interface supports the creation of tags within facets and supports the creation of tags without facets. Facette's tagging interface also enables the creation of new facets if existing facets are not satisfactory.

2.3.1 Basic UI Components

Facette's tagging interface provides an intuitive visualization of faceted information. Faceted tags and facets are represented using elements contained within a table. In Facette's tagging interface, rows are used to represent facets. A facet's name is contained in the first column of its row. A facet's values are contained in the second column of its row. In Figure 2.3, the 'type of object' facet is the first facet. This facet contains two tags.

Tags are represented using elements contained within the second column of the 'faceted tags' table. A tag's membership within a facet is determined by the row in which a tag is contained. In the example below, the tags 'aggregator' and 'blog' belong to the 'type of object' facet.

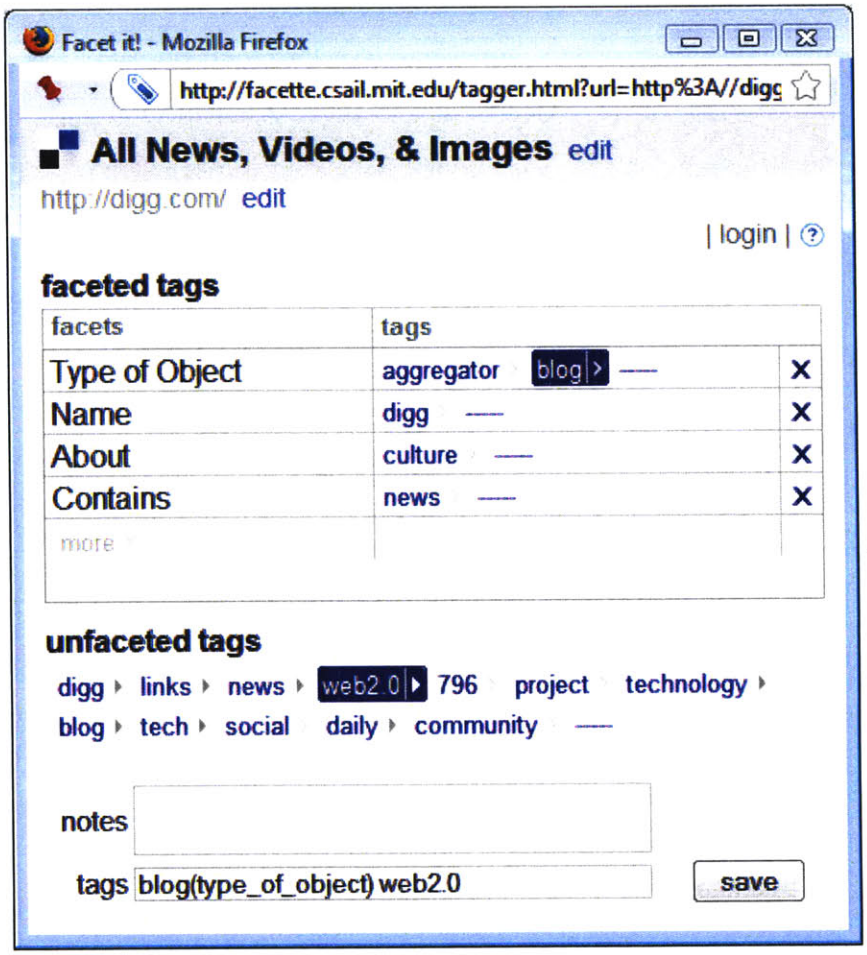

Figure 2.3: Tagging Interface

Note that the columns of the 'faceted tags' table are labeled 'facets' and 'tags'. These labels are provided to help reinforce row-based visualization of facet membership.

For tags without facets, an unbounded row is provided. Tags within the 'unfaceted tags' section of the interface resemble tags within the Delicious' tagging interface. These tags have no bounding containers and are associated with no facets. Through use of Facette's 'faceted tags' table and 'unfaceted tags' section, Facette's tagging interface is able to represent both faceted tags and unfaceted tags.

2.3.2 Tag Selection

Delicious' tagging interface distinguishes between a tag being selected and a tag being present. Selected tags are represented using white words embedded within blue boxes. Unselected tags are represented using un-stylized blue text. The clicking of these tags toggles their selection states.

For consistency with Delicious' tagging interface, identical selection behavior is used in Facette's tagging interface. Selected tags in Facette are represented using blue boxes. Unselected tags are represented using blue text. In Figure 2.3, the tags 'blog' and 'web2.0' are selected. Tag selection is toggle-able through clicking.

2.3.3 Tag Creation

Tags can be added to Facette's tagging interface through various input mechanisms. An input box that appears upon clicking the '-----' button allows for the creation of tags with arbitrary text. Facette's input box UI is shown in Figure 2.4.

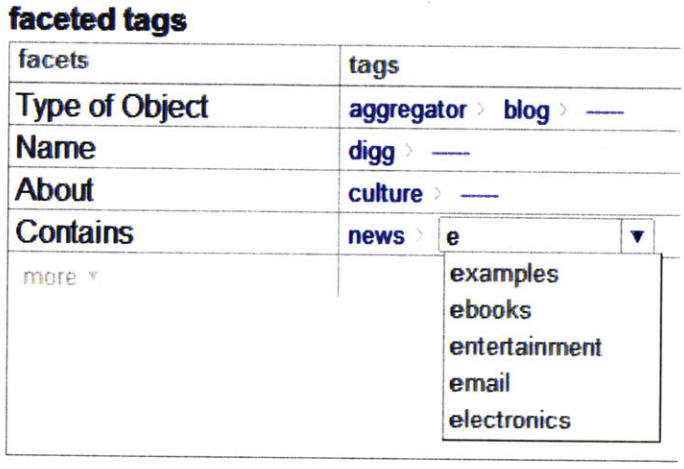

Figure 2.4: Input Box for Tag Creation

This input box appears after clicking the text '-----'. After the input box is shown, tag text can be typed. Pressing of the enter key or defocusing of the input box triggers the creation of a new tag using the text present within the input box. These input boxes allow for the creation of both faceted tags and unfaceted tags. A '-----' button is shown within each rows of Facette's tagging

interface. New tags are placed within the containing rows of the input boxes that were used to create them.

Within each input box, auto-completion is offered to assist the entering of text and to homogenize tag vocabulary. As text is typed, popular tag options are listed. For faceted tags, suggestions are derived using facet-specific popularity. For example, within the 'contains' facet, tags that would be suggested include 'art', 'humor', and 'news'. For unfaceted tags, a general list of popular tags from Delicious is used for auto-completion.

Faceted tags can be created through use of a second input mechanism by clicking upon menu items within tag-specific context menus. Each tag has a down arrow that activates of the display of a context menu. For tags that may belong within facets, recommended facets are shown. If a recommended facet is clicked, that tag is automatically copied to that facet. The follow figure shows a screenshot of an example tag context menu.

Figure 2.5: Context Menu for Tags

In this example, if the 'name' context menu item were to be click, the tag 'google' would automatically be created within the 'name' facet.

2.3.4 Facets

Table rows enable use of facets. To assist the use of facets, Facette shows a preliminary set of recommended rows. The following preliminary facets are included: 'type of object', 'name', 'about', 'contains'.

The creation of new facets is enabled through use of a facet input box that is shown after the 'more' button is clicked. The behavior of this input box is similar to that of Facette's tag input boxes. The following figure shows a screenshot of Facette's facet creation widget:

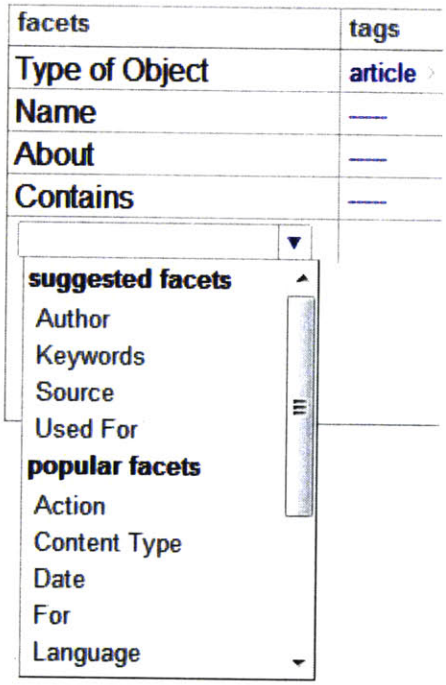

Figure 2.6: Input Box for Facet Creation

As with tags, auto-completion is used to assist the creation of facets. The generation of suggestions for facets is done through aggregation of previously submitted values. Popular facets are shown within a section of the list titled 'popular facets'. A second section is titled 'suggested facets'. This section was added to improve the usability of Facette's tagging interface during the early stages of development when little user data was available.

2.4 Collaborative Faceting

Facette uses collaborative faceting to assist the process of faceted tagging. Several auto-complete interfaces were described in the overview of Facette's tagging interface. In this section, these auto-complete interfaces are described in more detail.

2.4.1 Faceting Tags without Facets

Suggestions generated using collaborative faceting are used to help classify tags that have not yet been added to facets. In Facette's interface, Delicious's API is used to fetch relevant tags. These tags do not contain faceted information and must be manually faceted by the user. The following screenshot contains example tags that have been fetched using Delicious's API.

Figure 2.7: Example Unfaceted Tags

For the web page www.reddit.com, Delicious provides the tags 'bookmarks', 'social', and 'web2.0'. Many of these tags likely can be placed within facets; unfortunately, the challenge of choosing appropriate facets potentially makes such task difficult.

To assist the faceting of these tags, Facette provides facet suggestions for tags that have previously been placed within facets. A darkened gray arrow is placed adjacent to tags with existing facet suggestions. When this arrow is clicked, a context menu with facet suggestions appears. The following screenshot contains the context menu for the tag 'bookmarks'.

Figure 2.8: Facet Suggestions Dropdown for Tags

Because the tag 'bookmarks' has most frequently been placed within the 'Contains' facet, this facet is the first facet listed in the context menu. The two other facets have also been used to classify the tag 'bookmarks'. The algorithm used to generate these suggestions will be described in more detail later in this chapter.

2.4.2 Suggestions for New Facets

Suggestions generated using collaborative faceting are also used to assist the creation of new facets. Facette provides the user with a list of popular facets during the facet creation process.
As soon as the facet creation textbox has been activated, a dropdown containing suggestions appears. The following screenshot contains an example list of popular facets.

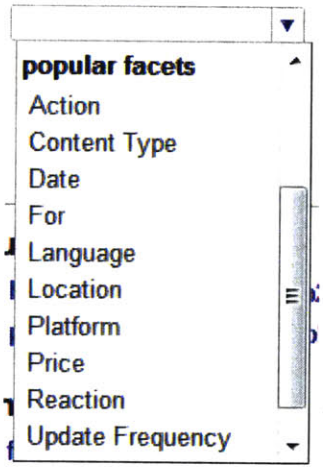

Figure 2.9: Facet Suggestion Dropdown for New Facets

This list contains several popular facets that are likely to help the user choose which new facets to include. Previously used facets appropriately classify a large percentage of the most commonly used tags on Delicious. Data supporting this claim is shown in Chapter 3 of this thesis. To create a list of popular facets, Facette's web service aggregates previously submitted information. The details of this process will be described in the algorithms section of the chapter.

2.4.3 Page-specific Suggestions

Finally, collaborative faceting is used to suggest faceted tags for pages that have previously been bookmarked using Facette. For pages that have previously been bookmarked using Facette, Facette's tagging interface is pre-populated with suggested faceted tags. The following screenshot contains example suggestions.

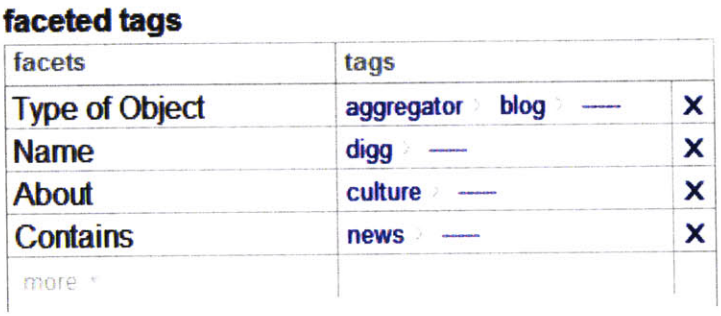

Figure 2.10: Example Suggested Faceted Tags

These suggestions are for the page www.digg.com. The faceted tags 'aggregator' and 'blog' have been suggested because users have previously selected these tags when bookmarking this page. The algorithm used to generate these suggestions will be described in the upcoming algorithms section of this chapter.

2.4.4 Suggestion Algorithms

The algorithms used to generate suggestions require the processing of faceted information and are designed to minimize implementation difficulty. Suggested facets for tags that have not been placed within facets are generated by calculating the facet histograms of common tags. For example, when generating facet suggestions for the tag 'audio', the following information is used: the tag 'audio' has been placed within the 'Type of Object' facet 3 times. This tag has been placed within the 'Contains' facet 3 times and has been placed with the 'About' facet once. Given such information, my facet suggestion algorithm would list the 'Type of Object' facet and the 'Contains' facet as suggested facets. The 'About' facet would not be suggested because it has only been used once. If a tag were used within more than 3 facets, only the most popular 3 facets would be recommended.

The algorithm used for generating suggested facets for dynamically created facets is equally simple. For these suggestions, Facette returns a list of Facette's most popular facets. To be a suggested facet, a facet must be used more than once. Thus far, 30 facets in Facette have been used more than once. More usage details are included in Chapter 3 of this thesis.

Finally, for creating suggested faceted tags, Facette calculates page specific tag histograms. For each page that has been bookmarked, Facette generates a tag histogram using faceted information contained within the last 10 submissions for a page. The suggested faceted tags shown in Figure 2.4: Input Box for Tag Creation indicate the previous users have used the 'aggregator' and 'blog' tags within the 'Type of Object' facet and have used the 'digg' tag within the 'Name' facet. Using Facette, users have generated faceted tag suggestion for thousands of bookmarks. More usage details are included in Chapter 3 of this thesis.

In review, collaborative faceting is used to generate suggestions for Facette's tagging interface. Collaborative faceting is used to generate suggestions for unfaceted tags, for dynamically created facets, and for previously bookmarked pages. For ease of implementation, simple algorithms are used to generate suggestions. All suggestions are generated through some variation of metadata aggregation.

2.5 System Architecture

For ease of implementation, Facette is designed as an augmentation of a popular online bookmarking service called Delicious. Delicious provides an immense amount of infrastructure necessary for the implementation of a custom tagging service.

Facette's tagging interface is designed to be compatible with Delicious service. Facette's faceted metadata is stored as an augmentation of Delicious tag syntax. When faceted metadata is created, the metadata is flattened for compatibility with Delicious. This flattened metadata is saved to Delicious using Delicious' API. To circumvent browser security, a browser plugin is required for the saving of data to Delicious.

Facette's browsing interface is designed to use data stored within a user's Delicious account. Facette's browsing interface uses Delicious' API to extract data. Faceted tags are directly loaded into the browsing interface. Unfaceted tags are preprocessed before being loaded. To improve the usefulness of Facette's browsing interface for people who have not faceted their own tags, unfaceted tags with facet suggestions are automatically faceted during the preprocessing of tag metadata. A caching system is built into Facette's browser plugin to reduce Facette's dependencies on data transfers.

An open-source faceted browsing framework called Exhibit is used to simplify the implementation of Facette's browsing interface. A specific instance of Exhibit is used by Facette to enable faceted navigation, tag navigation, and search of bookmarked information.

2.5.1 Delicious

Delicious [17] provides fundamental infrastructure necessary for implementing a custom tagging service. Delicious implements credential management for user accounts. To protect bookmarked information, a username and password is required for access to Delicious' service. After access is granted, Delicious provides online storage of bookmarked information. Delicious enables the saving of tag metadata within a structure that indexes information by URL. Finally, Delicious provides an API that supports the implementation of custom tagging interfaces and supports the extraction of user data. Delicious' API may be used to add tag metadata to a person's account. Delicious' API may also be used to retrieve saved information.

2.5.2 Tagging Interface

Facette's tagging interface is designed to use Delicious' API. Faceted metadata that has been generated using Facette's tagging interface is stored as an augmentation of Delicious tag syntax. When faceted metadata is created, the metadata is 'flattened' for compatibility with Delicious. Faceted metadata is flattened through use of faceted tag syntax. The following figure contains an example of faceted tag syntax.

tool(type_of_object) scribd(name) books(contains)

Figure 2.11: Faceted Tag Syntax

Faceted tag syntax enables embedding of faceted information within traditional tags. This syntax requires that a tag's facet be placed within parentheses and then appended to the end of a tag. Because our faceted tag syntax results in the generation of traditional tags, use of faceted tags syntax prevents faceted information from complicating metadata storage. As an added bonus, use of this syntax also preserves the natural alphabetical sorting of tags within Delicious if faceted tags were to be viewed using Delicious.

Flattened faceted information is saved to Delicious using Delicious' API. When the submit button is clicked in Facette's tagging interface, flattened faceted information is generated. A browser plugin is used assist the saving of this information to a person's Delicious account. The follow system diagram illustrates the interaction of Facette's tagging interface with Delicious.

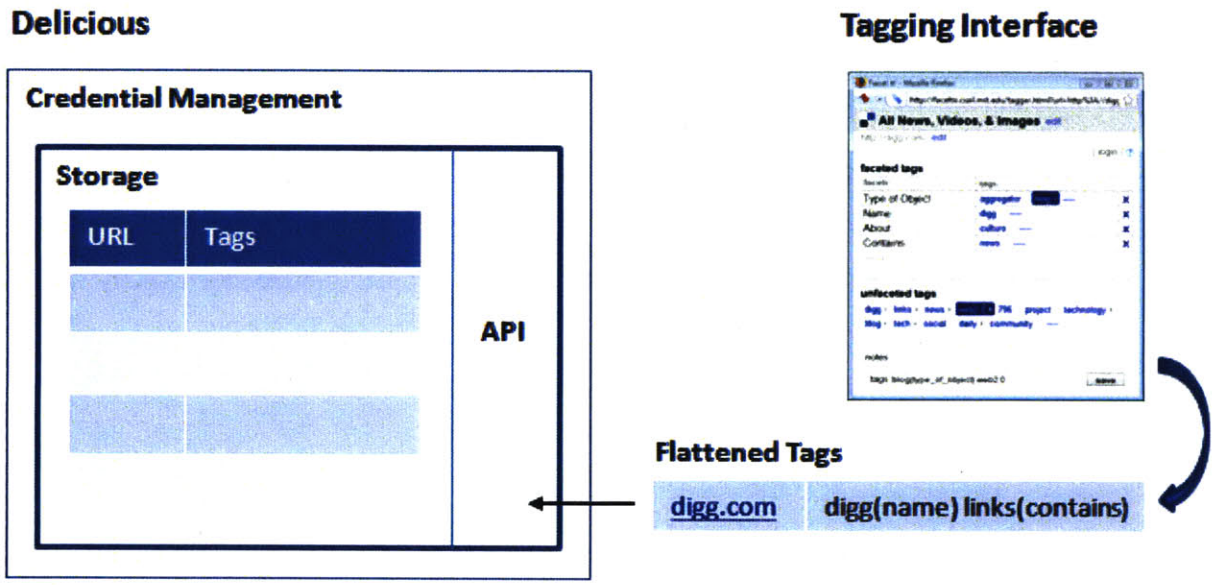

Figure 2.12: System Diagram for Tagging Interface

In review, Facette's tagging interface is used to generate faceted metadata. This metadata is flattened using faceted tag syntax. A browser plugin assists the saving of flatten information to a person's Delicious account using Delicious's API. When necessary, the user is prompted for account credentials.

2.5.3 Browsing Interface

Facette's browsing interface is designed to use data from a user's Delicious account. Facette's browsing interface is capable of parsing and augmenting Delicious tag metadata. Tags with imbedded faceted information are unflattened before they are loaded into Facette's browsing interface. Tags without imbedded faceted information are preprocessed before being loaded. Facette tries to augment unfaceted information by automatically placing tags within facets. If a facet suggestion exists for a particular tag, that tag is faceted before being entered into the system. This feature was added to improve the usefulness of Facette's browsing interface for users without faceted information. This feature can be disabled if it becomes troublesome.

To minimize the downloading of user data, Facette's browser plugin has a built-in caching mechanism. Data from Delicious's API is requested only if new bookmarks have been added to the user's collection since data was last downloaded.

In review, Facette's browsing interface is able to reconstruct faceted information from tag metadata extracted from Delicious. When used in combination with Facette's tagging interface, Facette's browsing interface is able to provide full use of concepts from faceted tagging.

Chapter 3
Evaluation

Since its public debut on February 16th, 2009, Facette has successfully assisted the use of faceted tagging. Facette has been used by several hundred Delicious users. Over 3,000 bookmarks have been created using Facette's tagging interface. Of people who have tried Facette, about a dozen have adopted Facette as their primary means of creating tags. A subset of these users also regularly use Facette's browsing interface.

Both faceted tags and unfaceted tags are created using Facette's tagging interface. Users create an average of 4.9 tags for each bookmark added using Facette. 55% of these tags are left unfaceted. The features within Facette's tagging interface have not discouraged the use of unfaceted tagging. The average number of unfaceted tags created using Facette's tagging interface matches the average number of unfaceted tags created using Delicious' tagging interface.

Several of Delicious's most popular tags have consistently been placed within facets using Facette's tagging interface. Amongst tags that are frequently placed within facets, tags that represent objects and web page topic are commonly represented. In addition, several of Delicious' most popular tags have consistently been left unfaceted. Users have abstained from faceting several tags resembling web buzzwords.

Facette's most frequently used facets predominantly belong to a predefined list of suggested facets; however, the enabling of free faceted tagging has yielded moderate use of several dynamically created facets. The 'For' facet and the 'Update Frequency' facet are examples of facets that have organically gained attention.

Over the past many months of use, enough faceted information has been collected to generate suggestions for several thousand pages, facets, and tags. Tag suggestions exist for roughly 3,700 web pages. Facet suggestions exist for roughly 760 tags. Finally, when dynamically adding new facets, a user is provided with a list of several dozen existing facets.

3.1 User Adoption

Several Delicious users have adopted Facette's tagging and browsing interfaces. Since February 16th, 2009, Delicious users have created over 3,000 bookmarks and have made an average of 4.9 tags per bookmark. Since Facette's launch, about a dozen users have been shown to consistently use Facette's tagging interface. A similar number of users also regularly visited Facette's browsing interface.

3.1.1 Active Tagging

55 people have used Facette's tagging interface to bookmark more than 10 web pages. About a dozen of these users have adopted Facette as their primary method of tagging. Amongst users who first started using Facette in either February or March, 14 users are still actively using Facette for bookmarking web pages. The following table contains Facette usage statistics for Facette's 14 most active users.

User ID	# of Bookmarks Created	Date Joined	% of Bookmarks with Faceted Tags	Monthly Visits to Browsing Interface
1	123 bookmarks	2/18/09	98%	0 visits/month
2	67 bookmarks	2/18/09	5%	0.5 visits/month
3	66 bookmarks	2/19/09	100%	7 visits/month
4	57 bookmarks	3/1/09	100%	1 visit/month
5	47 bookmarks	3/16/09	100%	1 visit/month
6	37 bookmarks	3/6/09	97%	3 visits/month
7	37 bookmarks	3/21/09	100%	4.5 visits/month
8	34 bookmarks	2/17/09	41%	0 visits/month
9	31 bookmarks	2/19/09	94%	0.5 visits/month
10	29 bookmarks	2/18/09	0%	6.5 visits/month
11	28 bookmarks	2/20/09	0%	3 visits/month
12	22 bookmarks	3/3/09	40%	3 visits/month
13	14 bookmarks	2/19/09	100%	2.5 visits/month
14	11 bookmarks	2/22/09	92%	2 visits/month

Table 3.1: Usage Statistics for Active Users

Facette's most active users consistently use Facette's tagging interface for bookmarking. Access data from Facette's server show frequent use of Facette by these users. Analysis of the users'

Delicious bookmarks reveals that a high percentage of the users' bookmarks contain faceted information. 9 of 14 frequent users show near 100% use of faceted information. Another two users show moderate use of faceted information. Oddly, 3 of 14 frequent users seem to neglect Facette's faceted tagging features despite the presence of such features.

Facette's most active users tend to begin use of Facette with the perceived intent of adopting faceted tagging. When beginning use of Facette, many users use Facette's tagging interface to retag several existing bookmarks within their Delicious accounts. The following figure shows example usage statistics of Facette's tagging interface for the first few days of an active user's account.

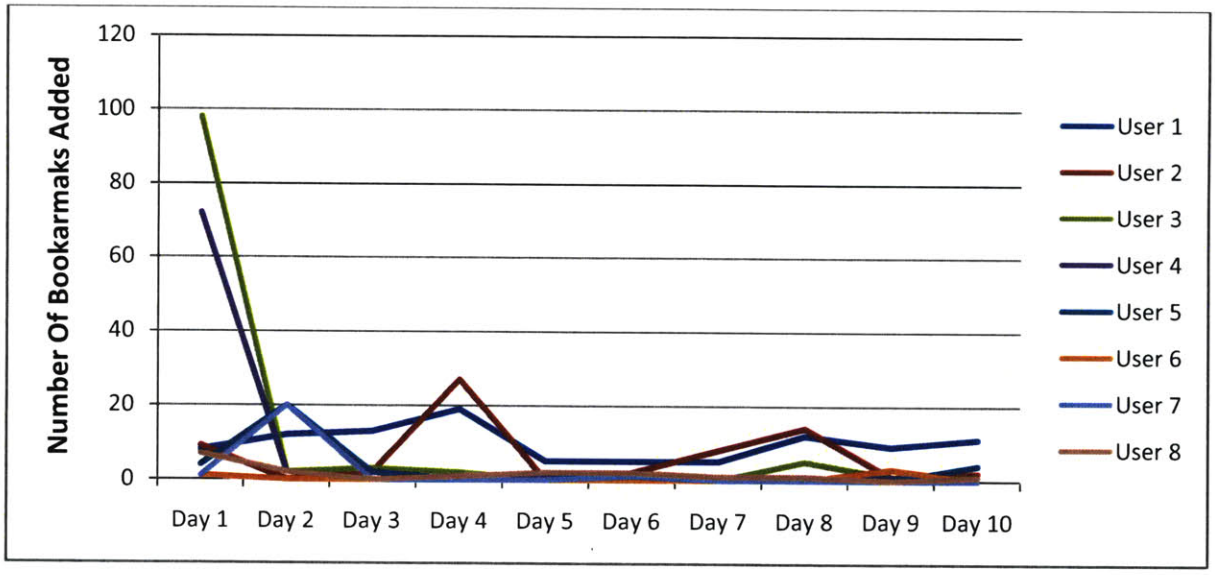

Figure 3.1: Usage Statistics for Tagging

After the first day or second day of exposure to Facette, many users retag about 20 or more existing bookmarks. After an initial flurry of retagging, a typical active user submits between zero and two bookmarks a day.

Of Facette's earliest adopters, a handful of previously active users have discontinued use of Facette's tagging interface. These previously active users began use of Facette with the retagging of several dozen bookmarks. Unfortunately, after a month of consistent use, these users seemed to have deliberately discontinued use of Facette's tagging interface. One hypothesis that

explains such discontinuation is that Facette's tagging interface does not offer some features included with Delicious's tagging interface. For future work, further investigation of this issue is suggested.

3.1.2 Active Browsing

Facette's browsing interface receives regular attention from a subset of active Facette users. Most of Facette's regular users seem to value Facette's browsing interface. Statistics for average monthly visits to Facette's browsing interface are listed in Table 3.1. 8 of Facette's 14 most active users visit Facette's browsing interface an average of two or more times a month. Although users use Facette's browsing interface substantially fewer times than they use Facette's tagging interface, a biweekly visit to Facette's browsing interface does suggest some level of dependence on the browsing of collected faceted information.

3.2 Tag Creation

Both faceted tags and unfaceted tags are created using Facette's tagging interface. For each bookmark that is tagged using Facette's tagging interface, users create an average of 4.9 tags. 45% of these tags are placed within facets. The remaining tags are left unfaceted.

When bookmarking pages, users often use a mix of faceted tags and unfaceted tags. 36% of bookmarked pages are labeled using both types of tags. Only 42% of bookmarks are labeled exclusively using faceted information.

Finally, the features within Facette's tagging interface have not discouraged the use of unfaceted tagging. The average number of unfaceted tags created using Facette's tagging interface matches the average number of unfaceted tags created using Delicious's interface. A study of Delicious has shown that Delicious user's create an average of 2.51 tags per bookmark [18]. Similarly, Facette's users create an average of 2.7 unfaceted tags per bookmark.

3.2.1 Creation within Facets

Several of Delicious' most popular tags have consistently been placed within facets using Facette's tagging interface. The following table contains faceting statistics for Delicious's 30 most commonly used tags.

Tag	# of Occurrences	% Placed within Facets
blog	407	86%
blogs	48	25%
news	139	54%
web	142	9%
design	138	44%
technology	141	60%
web2.0	199	13%
daily	28	54%
reference	575	86%
software	174	36%
internet	50	18%
media	44	9%
tech	20	15%
webdesign	48	31%
culture	38	47%
art	65	74%
programming	103	77%
tools	271	41%
inspiration	70	66%
business	107	44%
development	51	49%
community	29	7%
fun	28	14%
cool	18	50%
blogging	10	20%
politics	47	55%
music	78	46%
humor	50	66%
howto	127	37%
tips	179	75%

Table 3.2: Tag Faceting Frequencies

The third column of this table displays the frequency with which tags are placed within facets. For example, this table shows that the tag 'blog' is placed within a facet during 86% of tagging instances. In the other 14% of instances, the tag 'blog' is selected but is not faceted. Of Facette's 30 mostly commonly used tags, 8 tags are placed within facets during more than 65% of tagging instances.

Amongst Facette's most frequently faceted tags, several tags represent webpage object types. The two mostly frequently faceted tags shown within Table 3.2: Tag Faceting Frequencies describe a page's object type. Both these tags are placed within facets during 86% of tagging instances. The following table contains more examples of tags that describe object type and are frequently placed within facets.

Tag	# of Occurrences	% Placed within Facets
Article	662	100%
Reference	492	86%
Blog	350	86%
Homepage	336	100%
Video	127	64%
Website	77	94%
Tutorial	82	69%
Service	54	90%
List	36	84%

Table 3.3: 'Type of Object' Tag Faceting Frequencies

The tags in Table 3.3: 'Type of Object' Tag Faceting Frequencies represent the 9 tags that are most commonly used within the 'Type of Object' facet. Most of these tags possess a faceting frequency of greater than 85%.

Tags that describe webpage topics are also commonly placed within facets. From Table 3.2: Tag Faceting Frequencies the tags 'technology', 'programming', and 'politics' each describe webpage topics and have faceting frequencies of greater than 50%. More example tags can be found within the 'About' facet. The majority of tags that appear most frequently within the 'About' facet have faceting frequencies above 40%.

Finally, tags describing collections of items are commonly placed within facets. The following tags from Table 3.2 describe collections of items: 'blogs', 'news', 'tools', and 'tips'. The tags 'art' and 'music' may also be included within this set. With the exception of the tag 'blogs', all of these tags have faceting freqencies of greater than 40%. More example tags can be found within the 'Contains' facet. Many tags that are commonly placed within the 'Contains' facet resemble plural nouns and have faceting frequencies of greater than 50%.

3.2.2 Unfaceted Tags

Within Facette's bookmarking service, there exist several tags that users refrain from faceting. Amongst infrequently faceted tags, web buzzwords are commonly represented. Of Facette's 25 most popular unfaceted tags, the following tags have faceting frequencies of less than 15%.

Tag	# of Occurrences	% Placed within Facets
web2.0	100	14%
twitter	104	11%
web	142	9%
mobile	69	3%
cloud	60	0%
social-networks	55	0%
social	115	1%
test	63	2%

Table 3.4: Faceting Frequencies for Popular Unfaceted Tags

The tags 'web2.0' and 'mobile', 'cloud', 'social', 'social-networks', and 'twitter' describe trending topics on the web. Together, these tags represent 75% of the tags listed within Table 3.4: Faceting Frequencies for Popular Unfaceted Tags.

In conclusion, both faceted tags and unfaceted tags are created through use of Facette's tagging interface. When bookmarking pages, most users select a mix of faceted tags and unfaceted tags. Amongst tags that are frequently placed within facets, tags representing objects and webpage topics are commonly represented. In contrast, amongst tags that are infrequently faceted, tags representing web buzzwords are commonly represented.

3.3 Facet Usage

Facette's most frequently used facets predominantly belong to a predefined list of suggested facets. The following facets are shown as suggestions within Facette's tagging interface: 'Type of Object', 'Name', 'About', 'Contains', 'Author', 'Source', 'Used For', and 'Keywords'. These 8 suggested facets account for 95% of all facets used during the tagging process. A full histogram of facet usage is shown below:

Facet	# of Uses	% of Uses
Type of Object	1632	31%
About	1621	30%
Contains	891	17%
Name	515	10%
Used For	209	3.9%
For	131	2.5%
Author	89	1.7%
Source	84	1.6%
Keywords	52	1.0%
Written In	37	0.7%
Language	25	0.5%
Location	17	0.3%
Update Frequency	15	0.3%
Platform	8	0.1%
Action	7	0.1%
Price	4	0.1%
Date	3	0.1%

Table 3.5: Histogram of Facet Usage

In this table, suggested facets are shown in red. The top three most frequently used facets account for near 80% of all facet use.

User-created facets are much more rarely selected; however, they are important for the faceting of certain types of common tags. The 'Language' facet is used for the faceting of tags that describe languages. Popular tags within this facet include the tags 'english' and 'french'.

The 'Written In' facet is used for the faceting of tags that describe programming languages used for website development. For example, the 'flash' tag has been placed within the

'Written In' facet for pages that have been implemented using Flash. Other common tags within this facet are 'java' and 'javascript'.

The 'For' facet is the most popular user-created facet. Tags within this facet describe what a website is for. Tags within this facet are similar to tags within the 'Used For' facet; however, they include a broader variety of tags. Example popular tags within this facet are 'dinner', 'firefox', and 'research'.

In review, most tags are placed within suggested facets. 95% of faceted tags are placed within one of 8 suggested facets. Nonetheless, user-created facets are important for the faceting of certain types of tags. The 'Language' facet is used for faceting language related tags. The 'Written In' facet is used for the faceting of tags representing programming languages. The 'for' facet is used for describing what a website is for.

3.4 Collaborative Faceting

Over the past many months of use, enough faceted information has been collected to generate suggestions for several thousand pages and several hundred tags. Faceted tag suggestions exist for roughly 3,700 pages. Most pages within Facette's have been tagged by only one user. Of the 3,700 pages for which Facette has faceted information, 3,600 of these pages have been tagged only once. Nonetheless, suggestions are available for all pages that have previously been tagged. Using Facette's suggestion algorithm, all submitted faceted information is processed for suggestion generation.

Enough faceted information has been collect to yield facet suggestions for many of the tags used within Facette's interface. For a tag to obtain a facet suggestion, Facette's users must independently place a tag within the same facet at least twice. For example, in order Facette to suggest that the tag 'photos' be placed within the 'Contains' facet, two users must independently place the 'photos' tag within the 'Contains' facet. Of the 3670 unique tags that have been created using Facette's tagging interface, 760 tags possess facet suggestions that have been generated through use of collaborative faceting.

Finally, when dynamically adding new facets, a user is provided with a list of several existing facets. Facette's suggestions are comprised of the facets listed within Table 3.5: Histogram of Facet Usage. In total, 30 facets have been created using Facette's facet creation textbox; however, to help eliminate noise, the least used facets have been hidden from Facette's users.

In review, the processing faceted information from Facette has yielded suggestions for several thousand pages and several hundred tags. Most pages bookmarked using Facette has suggested faceted tags. Many of Delicious's most popular tags have facet suggestions, and when creating new facets, Facette's users are shown several dozen existing facets.

Chapter 4
Conclusion

4.1 Future Work

Further evaluation is recommended for several untested aspects of Facette. Further evaluation could be conducted to learn why some users stop using Facette after an extended period of use. One hypothesis is that features within Facette's tagging interface are perceived as unhelpful after being used for bookmarking several websites. If such hypothesis were true, users who abandoned Facette's tagging interface might still value features within Facette's browsing interface. Preliminary evidence suggests that this hypothesis may be true. Some of Facette's users disregard the facet-related features within Facette's tagging interface despite their continued use of Facette's browsing interface. A user study would help tremendously with such evaluation.

Future work could also focus upon an evaluation of why people use Facette's browsing interface. Preliminary information is able yield some insight. Facette's users seem to enjoy having a browsing interface that unifies concepts from faceted navigation, tag navigation, and search. Facette's user may also enjoy the dynamic behavior that Facette's interface inherits from Exhibit. Again, a user study would help tremendously with collecting relevant information.

Finally, additional enhancements to Facette's browsing interface are recommended. Facette's browsing interface is capable of rendering unique views for items within a collection of bookmarks. For web pages containing photos, embedded photos could be displayed. For web pages associated with locations, a geographical map could be displayed. The navigation controls within Facette's interface are also capable of adaptive rendering. For faceted information relating to time, time-based selectors could be shown. For faceted information relating to color, a color selector could be shown. For faceted information relating to location, a location selector could be shown. Various other possibilities likely exist.

4.2 Contributions

In this thesis, we introduce the concepts of free faceted tagging and collaborative faceting. Free faceted tagging is a manual method of faceting that allows users to create both tags and facets during the process of tagging. Free faceted tagging helps resolve tag ambiguities introduced during the tagging process and encourages the creation of faceted information. Free faceted tagging allows users to dynamically create facets. In addition, free faceted tagging does not require that all tags be placed within facets.

Collaborative faceting describes the process of generating tag and facet suggestions based upon previously generated faceted information. Collaborative faceting is capable of generating suggestions for unfaceted tags, dynamically created facets, and web pages.

To test these novel concepts, this thesis introduces a proof-of-concept system called Facette that implements free faceted tagging and collaborative faceting. To browse bookmarked information, Facette uses a browsing interface that displays both tags and facets. To create faceted information, Facette uses a custom tagging interface that incorporates concepts from free faceted tagging and collaborative faceting. For ease of implementation, Facette is designed as an implementation of Delicious tag-based bookmarking service.

Deployment of Facette has yielded successful results. Since its public debut on February 16th, 2009, Facette has been used by several hundred Delicious users. Facette's tagging interface has led to the bookmarking of thousands of pages with thousand of faceted tags. Both faceted tags and unfaceted tags have been created using Facette's tagging interface. Enough faceted information has been collected to generate useful suggestions for thousands of pages, for hundreds of tags, and for dynamically created facets.

Bibliography

[1] *Clustering versus faceted categories for information exploration.* **Hearst, Marti A.** s.l. : ACM, 2006, Commun. ACM, Vol. 49, pp. 59--61. 0001-0782.

[2] **Huynh, D F.** *User Interfaces Supporting Casual Data-Centric Interactions on the Web.* 2007. phdthesis.

[3] *The structure of collaborative tagging systems.* **Golder, S and Huberman, B A.** 2005, Arxiv preprint cs.DL/0508082.

[4] *tagging, communities, vocabulary, evolution.* **Sen, Shilad, et al.** s.l. : ACM, 2006. pp. 181--190.

[5] *Improving tag-clouds as visual information retrieval interfaces.* **Hassan-Montero, Y and Herrero-Solana, V.** 2006. pp. 25--28.

[6] *Tag Clouds: Data Analysis Tool or Social Signaller?* **Hearst, M A and Rosner, D.** 2008. p. 160.

[7] *Folksonomies-cooperative classification and communication through shared metadata.* **Mathes, A.** 2004.

[8] *Faceted metadata for image search and browsing.* **Yee, Ka-Ping, et al.** s.l. : ACM, 2003. pp. 401--408.

[9] *The perfect search engine is not enough: a study of orienteering behavior in directed search.* **Teevan, Jaime, et al.** s.l. : ACM, 2004. pp. 415--422.

[10] *Flexible search and navigation using faceted metadata.* **English, J, et al.** 2002, Unpublished manuscript.

[11] *Thoughts on Social Tagging.* **Hearst, Marti.** Berkeley, CA : UC Berkeley School of Information, 2007.

[12] *Facetag: Integrating Bottom-up and Top-down. Classification in a Social Tagging System.* **Quintarelli, E, Resmini, A and Rosati, L.** 2007, IA Summit, Vol. 2007.

[13] *Nearly-automated metadata hierarchy creation.* **Stoica, E and Hearst, M.** 2004. pp. 117--120.

[14] *Inducing ontology from flickr tags.* **Schmitz, P.** 2006.

[15] *Automated Tag Clustering: Improving search and exploration in the tag space.* **Begelman, G, Keller, P and Smadja, F.** 2006.

[16] *Structured vs. unstructured tagging--A case study.* **Bar-Ilan, J, et al.** 2006.

[17] Delicious. *http://delicious.com/.*

[18] *Understanding del.icio.us Tag Choice Using Simulations.* **Wash, R and Rader, E.**

[19] *Exploring Social Annotations for the Semantic Web.* **Wu, Xian, Zhang, Lei and Yu, Yong.** Edinburgh, Scotland : ACM, 2006. 1-59593-323-9.

[20] *Folksonomies: tidying up tags?* **Guy, M.** s.l. : University of Bath, 2006, D-Lib Magazine, Vol. 12.

[21] *Next generation web search: Setting our sites.* **Hearst, M A.** Bulletin of the Technical Committee on, p. 38.

[22] *Social Semantic Bookmarking.* **Braun, S, Zacharias, V and Happel, H J.** 2008. pp. 62--73.