Collaborative Categorization on the Web: Approach, Prototype, and Experience Report

François Bry and Holger Wagner

Institute for Computer Science, University of Munich
Oettingenstr. 67D-81538 München, Germany
http://www.pms.informatik.uni-muenchen.de/
September 2003

Abstract

Collaborative categorization is an emerging direction for research and innovative applications. Arguably, collaborative categorization on the Web is an especially promising emerging form of collaborative Web systems because of both, the widespread use of the conventional Web and the emergence of the Semantic Web providing with more semantic information on Web data. This paper discusses this issue and proposes two approaches: collaborative categorization via category merging and collaborative categorization proper. The main advantage of the first approach is that it can be rather easily realized and implemented using existing systems such as Web browsers and mail clients. A prototype system for collaborative Web usage that uses category merging for collaborative categorization is described and the results of field experiments using it are reported. The second approach, called collaborative categorization proper, however, is more general and scales better. The data structure and user interface aspects of an approach to collaborative categorization proper are discussed.

1 Introduction

The World Wide Web is a rich and complex space for retrieving all kinds of information in academic and commercial contexts. Many people are already collaborating in the effort to use this medium efficiently, but doing so without dedicated technical support. Instead, people are sending *URI*s pointing to documents they find interesting via email, with annotations that are lost when the email is deleted. Bookmark collections are manually converted to Web pages and uploaded to the Web - but peers must still be informed about the location of such pages and maintaining them is cumbersome. Some Web pages offer guest books or discussion forums, that are then used by an emerging community of people interested in the contents of such pages - but lacking an integrated system, their users have to cope with different user interfaces, no interaction between the discussion forums and generally no unified experience across different platforms.

The present paper is based on the ongoing development of an integrated approach to collaborative Web usage - supporting these forms of collaboration with a software system that integrates features existing distributed among various applications into a consistent concept. In *[Wagner2002]*, a terminology has been laid out, defining the terms relevant to the subject matter. Concepts relevant to an integrated approach to

collaborative Web usage have been introduced there: collaboration, communities, Web navigation, communication, categorization and privacy and security issues. This paper refers to that terminology.

2 Relevant Concepts and Approaches

An effort to clarify the terminology for the broad area of the World-Wide Web has been summarized in *[W3CWCA]*. *[Cheung]* defines a *Web tool* as a software tool that helps users to retrieve, locate and manage Web documents and provides a classification for such Web tools into five levels.

Work dealing with the creation and integration of user interfaces for revisitation and annotations tools includes *[Barret]*, *[Cockburn99a]*, *[Cockburn99b]*, *[Hascoet1999]*, *[Hascoet]*, *[Kaasten]*, *[Koch]*, *[Laurent]*, *[Li]*, and *[Tauscher]*.

There have been various approaches to annotating the WWW. One major design issue with annotation systems is how the annotations are gathered, stored and presented. There are generally two classes of systems: systems that require software installation or configuration changes on the client-side (e.g., *[Kahan]*, *[Laurent]*, *[Marais]*), and systems that use standard internet technology like JavaScript to embed the functionality in standard *Web browser*s (e.g., *[Koch]*).

The need for a tool for collaborative *Web usage* is illustrated in *[Twidale]*. That paper draws from findings on how conventional libraries are used by students - namely, often in a collaborative manner - and these findings can be transferred to World-Wide Web usage. One interesting idea in this work is that not only information, but also people are considered an important thing one can search for:

> "We believe that browsing for people, their electronic representations or representations of their activities, is a neglected and important area." (*[Twidale]*)

This can be addressed by unifying entities like *Web resources*, *information producers*, and *information consumers* into a single type of general entity and making the original type of the entity (e.g. *Web resource*) a category under which this general entity can be found. Therefore, a search concerning a specific topic may have all of the named entities in its result set if the user wishes so.

[Twidale] also introduces some interesting terms concerning (collaborative) browsing behaviour. Among others: *consulting* (asking other people for help) and to *bibble* (using other searchers published results). Since up to now, there is no proper mechanism for storing and sharing the information gathered from individual's searches, "the great majority of searches that are conducted fail to *bibble* properly" (*[Twidale]*) – a problem which is addressed in the present work by making individuals' bookmarks and the relevant categories available to communities.

[Marais] define *cooperative surfing* as activity of a community of users who cooperatively and asynchronously build up knowledge structures relevant to their group. A broad overview on collaborative Web usage is also given by *[Greenberg]*.

2.1 Privacy and Security

While the maximum protection of privacy may be an important criterion for many users (see *[Pitkow]*), this conflicts with the intention of making the Web more personal and supporting collaborative Web usage. Thus, a significant challenge for collaboration systems is balancing the protection of privacy with the rendering of personal information. One dimension of this is how much data is available about each user to which other users. *[Terveen]* suggested letting the users progressively reveal more about themselves, while they get to know the fellow users better (this is common practice with dating services).

In *[Bellotti]*, a very useful design framework is given, which is based on *control* and *feedback*. *User*s should be able to control what information about them becomes available to which other users and the users should be notified when information about them is being captured.

2.2 Collaborative Categorization

The concept of *collaborative categorization* is already used by other authors (for example, *[Lifantsev]*) and to promote the Semantic Web (for example by *[Allen]*) but doesn't seem to be well established. The focus of the present paper is finding ways to support the collaborative effort of creating and managing hierarchies of categories. We assume that once the categories themselves are handled in a collaborative manner, the categorization of actual items (e.g. *Web resources*) can be achieved in a natural, collaborative manner.

2.3 Further Relevant Concepts

An important concept relevant to categorization is that of ontologies, but the two concepts are somewhat orthogonal. *Ontologies* are created top-down, categorizations are created bottom-up: While ontologies must take into account any possible categories right from the start, categorizations start with the currently visible items and then are augmented when new items have to be categorized. Therefore, categorizations tend to grow in an evolutionary process and may become rather chaotic when there are no means to clean-up an existing structure. While this may also apply to growing ontologies, the problem is much more severe in categorizations. On the other hand, while ontologies require significant know-how in the domain they shall be applied to, categorizations can be created naively.

There are already existing systems for collaboratively creating ontologies, for example the Ontolingua Server. For a report on this tool, see *[Farquhar]*.

3 Categorization: An Essential Aspect of Collaboration on the Web

There are several areas of interest in the field of collaborative usage of the Web, where categorization plays a significant role or at least can be used to improve the efficiency of collaboration. The following areas are especially relevant to the research reported about in this paper:

- **Communities:** The most central concept of collaboration is a *community* of which the members are involved in collaboration. When a system for collaboration is targeted at a large user base, there will be a natural evolution of communities and keeping them organized well is not a trivial task. The goal of categorization of communities is that people with a specific interest find the community that best fits this interest. Multiple communities with the same center of interest should be avoided as well as single communities that are unspecific, resulting in a bad signal to noise ratio. If such general communities are needed, a good solution may be having *super-communities* that consist of multiple more specific communities. Hierarchical categorizations provide a natural means to manage such *super-communities*. From an abstract point of view, a community can be seen as a category for users.
- **Bookmarks:** Collections of pointers to Web resources where each such pointer can be put into one or more categories for later retrieval. Examples are hierarchically categorized directories like Yahoo and the typical bookmark collections stores in Web browsers. Collaborative usage of the Web can be achieved by making the individual bookmark collections accessible to peers.
- **Annotations:** Annotations are naturally categorized by the Web pages they are attached to. However, there may be multiple categories (i.e. types) of annotations attached to a single piece of information. These types may be motivated by different perspectives or goals under which the Web resources are used.
- **Communication:** While annotations are also a form of communication, the difference is that with annotations, an existing Web resource is enriched. The term *communication* is used in this context, when information that stands on its own is exchanged between people (even though it may refer to other pieces of information).
  There are several aspects of communication, where categorization plays a significant role. First, the channels of communication can be organized in categories. For example, communication within communities (categorization of communities' is inherited), personal communication, persistent vs. transient communication. However, an individual may prefer having an own categorization for all communication that he participates in (receives and sends). Then, this individual categorization may be aided by existing categorizations.

4 Making Collaborative Categorization a Central Aspect of Web Collaboration

The key thesis of this paper is that in the areas mentioned above, *collaborative categorization* will significantly improve the efficiency of collaboration. In fact, we assume that providing a means to effectively collaborate on building categorizations as an artefact in itself is the fundamental key to making collaborative usage of the World Wide Web common to a large user base.

Furthermore, from the findings of previous work we deduce that an integrated approach to collaborative Web usage can only be achieved if it is based on an integrated collaborative categorization system that does not separate the different means for collaboration (communities, annotations, bookmarks and communication) but instead unifies these means.

Hence, while in other approaches categorization usually implicitly plays a more or less important role, we pull collaborative categorization to the center of attention.

4.1 Application-Scenario: Web-Based Collaborative Software Development

In the process of software-development, one often has to make decisions on whether or not certain functionalities are better implemented in house, or whether the functionalities are already implemented by component developers and can therefore be bought and integrated into the system.

Usually, a project is implemented by a team - and during their work, they might encounter several components in question, storing the relevant Web pages describing these components as bookmarks. These bookmarks can be put into categories - for example:

- *application server*
- *billing*
- *user management*
- *content management system*

Even though many of the components that were found may not be useful for that particular project, they may be relevant for later projects, implemented possibly by a different team.

In a system that supports collaborative categorization, not only can the team members store their results in a categorization common to that team - but at a later point in time, another team could take several of the categories as a starting point for a new project.

For example, there may be a new project which does not require *billing* but does require *application server*, *user management* and a *content management system*. Furthermore, the component *content management system* might have to make available documents in a format suitable for printing, e.g. PDF. One team member could now add the existing categories to the new project (including the categories' contents that have been found in the previous project, e.g. bookmarks), and add a new, possibly empty category called *PDF generation*.

The longer such a system is in use, the more categories and actual items are available. However, as a system for collaborative categorization supports simplified views on the complex global categorization, it remains useful as long as someone makes a selection of the categories relevant to a particular project and/or team.

4.2 Application-Scenario: Literature Research

In the scientific community, there are many different fields that are connected in many different ways. A single team that has a specific research focus could use a collaborative system to collect and categorize information relevant to the work of that team. Obviously, this can improve the efficiency of that particular single team.

For example, a team that is focussed on collaborative usage on the Web may have a general category *collaboration* with several subcategories like *communication*,

annotations, *bookmarks* as well as *existing commercial systems* and *ongoing scientific projects*. If these categories and optionally the items stored within these categories (e.g. bookmarks to Web pages, articles and authors) are made public, this facilitates collaboration between teams. Not only can the focus of a team be determined by the categories they consider relevant to their work, but different teams with different foci can collaborate easily.

Imagine another team that is specifically interested in annotation systems. To them, *collaboration* is not an interesting general category, instead they use *annotations* as the most general category. Below that, they have one category *annotation systems* with several subcategories like *javascript based systems*, *centralized systems*, *distributed systems*, *server-based systems*, *client-based systems*, *proxy-based systems* and so on.

If the two teams decide to collaborate, they might share the *annotations* category even though for one of the teams, it's just one subcategory of *collaboration* and for the other team, it's the most general category. Furthermore, the first team may decide not to have the subcategories below *annotation systems*, but instead accumulate all items in all subcategories of *annotation systems* into that single category. While for the team focussed on *annotations*, the difference between *server-based systems* and *proxy-based systems* does play a significant role – the other team just needs to know the existing systems but does not care about the subcategories.

4.3 Application-Scenario: News Channels

Imagine a very large software company that needs to pay specific attention to their reputation in the public of different countries. For the heads of such a company, regular category-based press reviews can be a very helpful instrument for keeping informed of the public opinion. However, creating high-quality press reviews is not a trivial task. But such a task can be improved and simplified significantly by letting several individuals or teams collaborate in a fashion of division of labour.

For example, the categories may be created by the heads that need to use the press review. This may be a distributed team in which some of the members are interested in other categories than their team mates. Furthermore, some may have different priorities for the different categories. So the responsibility for creating and maintaining the categories both on the team level and individual level remains with the heads team.

Furthermore, there can be a team specialized to creating press reviews. Each team member may have different foci, e.g. *technical press* (concerning the technical quality of the company's products), *economic press* (concerning how the company's economic growth is seen in public) and *critical press* (concerning particularly critical articles – note that the same article may be put into several categories).

Assuming that the company has very open structures, all employees may be encouraged to add content to the system. However, to avoid degrading the quality of the press review, these employees get their own category where they can add interesting news items. The heads could access this category directly, however, to save time they may decide not to do so. However, members of the press review team may use the input from the employees and add it to the other categories, making the items more visible to the heads team.

4.4 Key Concepts of Collaborative Categorization

- **Inheritance:** There are many cases, where pieces of information (e.g. Web pages, annotations, messages) can inherit a categorization from their context. E.g. an annotation can be automatically be put into the same category as the Web page it belongs to - however, it may belong to other categories as well. Messages can be categorized in many ways from the context in which they were authored, e.g. community(ies) they were addressed to, Web pages that were previously viewed (assuming that there is a semantic connection between the page(s) viewed and the message).
- **Adaption to Single Users:** A global categorization system naturally becomes very complex, and it can be assumed that there is no global categorization that fits all individual needs. Therefore, it is crucial for such a system to support forms of adaption to individual users, hiding the complexity while keeping the full category tree in the background.
- **Adaption to Communities:** Communities of users may share concepts of how to categorize items in their particular field(s). A global categorization scheme in the background can be helpful a basis for community adaption. Such global categorization schemes can also be useful to compare how communities categorize and possibly also to exchange such information between communities – which might be seen as interesting approach to merge communities or community categorizations.

5 teamXweb: An Implementation of Collaborative Categorization by Category Merging

During the work on an approach to collaborative Web usage, we have developed a prototype system called teamXweb. The major idea behind this system was integrating various aspects of collaboration and making them accessible anywhere on the Internet. This was achieved by using a meta-browser architecture that implements an HTML-based Web browser on top of any existing Web browser. This meta-browser is implemented as a Web application that can be used within any Web browser that supports JavaScript and cookies.

In the prototype system, a simple form of Collaborative Categorization was implemented: *Category Merging*. While category merging does not provide the power of what we call *Collaborative Categorization Proper*, which will be introduced later in this paper, it is a simple and useful approach applicable to situations where complex categorizations are not needed and categorizations change less frequently.

In the following sections, the parts of teamXweb where Category Merging has been implemented are introduced.

5.1 Communication

Communication takes place in several areas of the system:

- Private Messages
- Community Messages

- Annotations
  - Domains - various scopes:
    - Personal
    - Community
    - Public
  - Pages - various scopes:
    - Personal
    - Community
    - Public

These different places can be seen as categories. Some of these categories have their own contexts, e.g. annotations belong to Web pages or Web sites and are usually accessible from these places. Therefore, these pages and sites can also be seen as categories. Furthermore, communities are such areas, which can also be seen as categories.

One special thing about teamXweb is that all these different areas are unified with a single interface so that all communication can be accessed from one view. This is seen as high-level category merging, as categories from various places are pulled together to be seen through a single interface.

This view is adapted to each user in several aspects:

- Only communities the user is a member of are shown for the community messages.
- Only those Web pages and Web sites in which annotations the user can read are available are shown to the user. For example, if there is an annotation for a particular Web page, visible to a particular community the user is not a member of, that annotation is obviously not shown. But even further, if there are no annotation to that Web page that the user could view - the whole Web page is not shown.

5.2 Bookmarks

With bookmarks, teamXweb implements the typical case of collaborative categorization via merging of categories. The user must select the scope, which can be either one of:

- personal bookmarks
- bookmarks of one of the communities the user is a member of

With personal bookmarks, only his personal bookmarks are accessible in the categorization he has created. However, when he selects a community, categories of all members of that particular community that have given the permission for other community members to view their bookmarks are merged.

As teamXweb only implements a flat categorization, this is very straightforward: all categories with different names are included in the view, containing the bookmarks they contain. If one or more categories with the same name exist for multiple users, the bookmarks contained in these categories are merged into a single category with that name.

That way, each user has his own categories - but by defining categories with the same name, users can collaborate on building the contents of that category.

While the approach with a flat categorization had been chosen due to time constraints during the implementation, merging categories also becomes very complex with hierarchical categorizations. There are several questions that cannot be solved easily, for example: If two categories with the same name exist in different sub trees - can these be merged at all? If they are merged, shall they be kept in both sub trees or a single one? If kept in a single sub tree - which sub tree is chosen? These problems make category merging useless in systems that need to scale with many users or systems that require hierarchical categorizations.

An approach that is comparatively simple to solve these issues with hierarchical categories is collaborative categorization proper, as introduced in section 7.

6 Experience Report

An experiment was pursued during 2002 in two phases of three months each. Selected communities of users, i.e. the attendants of three computer science courses and one ethnology course, all in all 4 courses and about 135 persons, have been given the previously described system. These communities have been assisted in learning how to use the system effectively. At the end of the experiment phase, each community member has been asked to answer questions by filling a Web-based questionnaire. About a fifth provided with extensive answers.

The questionnaire consisted of 10 questions. The first section, i.e. questions 1 to 5, was about usage of the system. In question 6, users were asked to give their opinion on eight features of the prototype. The remaining four questions were about how the system could be improved from a user's perspective. Note that the questions were not explicitly referring to collaborative categorization. Instead collaborative categorization aspects were implicitly addressed. Those results of this inquiry related to collaborative categorization are reported about here. For the sake of conciseness, the full statistical results are not given here, instead the salient aspect are reported about.

For each system feature, its practical importance and the quality of its realization could be evaluated using two separate scales for quality and importance. The scale for the quality had six values from `very good = 1` to `insufficient = 6`, the scale for the importance also had six values, ranging from `important = 1` to `useless = 6`.

As described in *section 5*, a meta-browser architecture has been chosen for the prototype, which was implemented as a Web application. The most important reason for choosing this approach was that no installation is required at the clients so that the system can be used easily anywhere on the Internet. The second part of question 6 asked how important users considered this feature and how they considered the quality of its implementation. As *figure 1* shows,

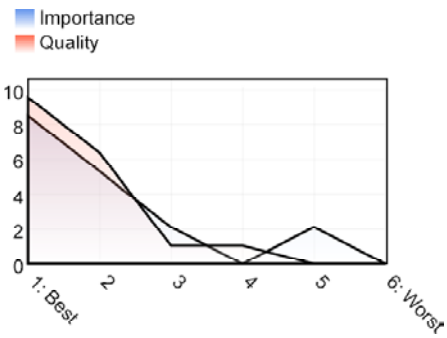

Figure 1: No Installation required

this has been considered both important and implemented well by most people. Therefore, building collaborative categorization systems based on this approach seems to be the right direction.

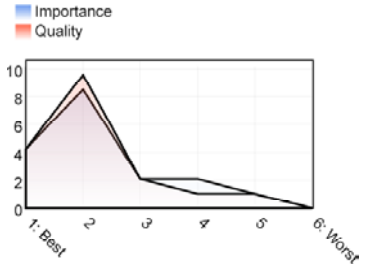

Figure 2: Sharing Bookmarks

The parts 3, 5 and 8 of question 6 dealt with the importance and quality of sharing bookmarks, sharing history, and annotations. As *figure 2* illustrates, sharing bookmarks has been considered equally important generally and implemented well in the teamXweb prototype by the users. The bookmark management in teamXweb was implemented with category merging and therefore this lets us conclude that not only is it important to provide means to share bookmarks, but also that collaborative categorization is a good approach for doing so.

Compared to bookmarks, the history of visited pages has been considered significantly less important (see *figure 3*) – a finding which is consistent with *[Catledge]* and *[Tauscher]*. Management of user's histories therefore does not seem to be an important area for collaborative Web usage in general and collaborative categorization in particular, even though it would make sense technically. One explanation for this result may be privacy concerns due to the passive recording of the history.

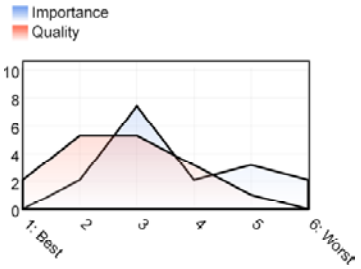

Figure 3: Sharing History

Finally, participants of the experiment were asked how important they considered annotations for individual Web pages and complete Web sites. The results are illustrated in *figure 4* and indicate that annotations systems should be accepted pretty well. Even though there is a "natural" categorization for annotations (due to the pages / sites they are attached to), other access methods based on collaborative categorization may significantly improve the usefulness of such annotations. In teamXweb, annotations were already integrated with other means of communication (within communities and between individuals) which seems to be a step in the right direction.

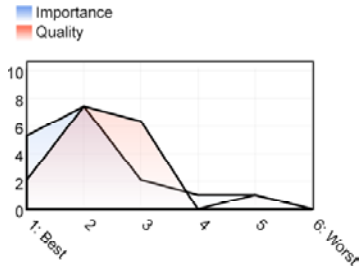

Figure 4: Annotations

7 An Approach to Collaborative Categorization Proper

7.1 Collaborative Categorization

The core idea behind Collaborative Categorization Proper is finding a mechanism to reduce the cognitive overload with too many categories for a single user while still

allowing every user to access every category if they need to. In particular, if different ways of categorization exist, it must be possible to have them coexist without confusing users.

7.2 Data-Model: Data Structures for Complex Categorizations

A solution to this problem is keeping a general tree (or possibly a general graph) that includes all categories of all users, and user specific views that include only a subset of the general tree. If a user adds a category, he needs to decide where it belongs in this general tree (that is, which is the parent category of the new category). Thus, consistency is assured as it is not possible to add categories without awareness of the context. Unfortunately, this does not prevent users from accidentally or intentionally putting categories where they do not belong. It may be useful to allow the same category having multiple parent categories breaking up the tree structure into a more general graph.

This general tree may grow to any complexity. For each user, however, there is a subset of this general large category tree, that only includes those categories that the user is interested in.

From the perspective of data structures, this means that in addition to the general tree, the user configurations need to be stored. The user specific tree configuration does not contain any categories - only whether or not a category from the main tree is displayed below a specific parent category. It is important, however, that it is not only possible to hide sub trees, but also paths. That way, it is possible to hide a complex structure while still providing access to a category anywhere within that structure.

Furthermore, the items stored under a given category *C* may be visible to its parent category *D* if *C* is hidden (inheritance of a category's items to its parent category). That way, all items are always available. However, hiding items with their categories must also be possible (this depends on the particular use case).

7.3 User-Interface: Aspects of Presentation and Interaction

Basically, the known user-interface elements for representing trees can be used to provide users access to the hierarchical categorization. Trees are so common in the current user-interface paradigms, so having a user-interface based on trees reduces the threshold for users to adapt to the new concept. Giving users access to a general graph structure may be useful for advanced users, though.

However, for gaining the full power of the underlying data structure, all users must be provided with additional actions: while in common trees, there are only three interesting actions: select an item, show children and hide children. For the multi-user-tree, the following additional actions are defined:

- **Hide Node and its Sub tree:** Hides a node (e.g., category) and all its ancestors from the view. If the edge is store (parent node and node), this can be used to hide a category in one place while keeping it below another category.
- **Show all Nodes:** This is the reverse action to the previous action. As hidden nodes cannot be selected, only their parent nodes can be selected and all their children be put back to view. Depending on the user interface implementation, a

more convenient method could be some sort of popup menu that allows selection of the desired node from a list of all hidden nodes of a given parent node.

- **Hide Node and inherit Children:** This can be used to flatten a deep tree. The node itself is hidden and all its children are shown as children of the node's parent node. However, the *weight* of the parent node may be significantly decreased this way. One consequence of the availability of this action is that the tree should be designed as deep as possible, with only few children for each node. While flattening such a tree is very easy with this feature, making a flat tree deeper is not possible.
- **Shortcut Path:** While in the previous action, all children of the hidden node are added to the node's parent node, this only adds a specific child. If repeatedly done, this allows removing complexity significantly, making a dense and rich tree very sparse.
- **Inherit / Hide Items of Invisible Categories:** A distinction must be made between categories and the items that belong to these categories. If this distinction is made, it is possible for each node that is hidden, to either hide or include the items belonging to the category represented by that node. With shortcut paths, it is necessary to allow inheriting items over multiple levels.

For the usage of such a categorization in a system for collaborative Web usage, it may be useful to allow users using the view configuration of their fellows. That way, they can profit of the effort others have put into customizing their categories.

Furthermore, default profiles could be provided with communities, so that each community could have its own categories while still being compatible with both its members' categories as well as the global set of categories (of which the member's categories are a subset). That way, a user could choose between the personal profile of himself or one of his fellows, and a public profile of one of the communities he is a member of.

8 Conclusion

Although not yet fully established, collaborative categorization is a very promising direction for research and innovative software systems. Arguably, collaborative categorization on the Web is especially promising because of both, the widespread use of the conventional Web and the emergence of the Semantic Web providing with more semantic information on Web data.

In this paper, collaborative categorization as such as well as collaborative categorization on the Web have been discussed. Essential aspects have been presented, especially the need to provide a consistent data structure that scales up well with a community's size, and the number of categorizes increase. Two approaches to collaborative categorization have been described and discussed: collaborative categorization via category merging and collaborative categorization proper. A prototype system for collaborative Web usage that employs category merging has been introduced. The relevant results of a field experiment with this prototype have been reported.

Conceiving, implementing, and testing a similar system for collaborative categorization proper seems to be a very promising direction for future research.

References

[Allen] Allen, J.: *Making a Semantic Web.* February 2001. Accessible at: http://www.netcrucible.com/semantic.html

[Barret] Barret, R., P. Maglio and D. Kellem: *How to Personalize the Web.* In Proc. CHI 97, ACM, 1997, pp. 75-82. Accessible at: http://www.acm.org/sigchi/chi97/proceedings/paper/rcb-wbi.htm

[Bellotti] Bellotti, V. A. Sellen: *Design for Privacy in Ubiquitous Computing Environments.* In G. de Michelis, C. Simone and K. Schmidt (Eds.) Proc. Third European Conference on Computer Supported Cooperative Work, (ECSCW '93), pp. 77-92. Kluwer, 1993.

[Catledge] Catledge, L.D. and J.E. Pitkow: *Characterizing browsing strategies in the world wide web.* Computer Networks and ISDN Systems, 27(6): 1065-1073, April 1995. Available at ResearchIndex (Citeseer): http://citeseer.nj.nec.com/catledge95characterizing.html

[Cheung] Cheung, D.W., B. Kao and J. Lee: *Discovering User Access Patterns on the World-Wide Web.* In Proceedings of the 1st Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD'97), February 1997. Available at ResearchIndex (Citeseer): http://citeseer.nj.nec.com/cheung97discovering.html

[Cockburn99a] Andy Cockburn, Saul Greenberg, Bruce McKenzie, Michael Jasonsmith and Shaun Kaasten: *WebView: A Graphical Aid for Revisiting Web Pages.* In Proceedings of the OZCHI'99 Australian Conference on Human Computer Interaction, Wagga Wagga Australia, November 1999. Available at ResearchIndex (Citeseer): http://citeseer.nj.nec.com/cockburn99webview.html

[Cockburn99b] Andy Cockburn and Saul Greenberg *Issues of Page Representation and Organisation in Web Browser's Revisitation Tools.* In Proceedings of the OZCHI'99 Australian Conference on Human Computer Interaction, Wagga Wagga Australia, November 1999. Accessible via: http://www.cpsc.ucalgary.ca/grouplab/papers/index.html

[Farquhar] Farquhar, A., Fikes, R., and Rice, J.: *The Ontolingua Server: a Tool for Collaborative Ontology Construction.* 1996. Available at ResearchIndex (Citeseer): http://citeseer.nj.nec.com/farquhar96ontolingua.html

[Greenberg] Greenberg, S.: *Collaborative Interfaces for the Web.* In C. Forsythe, E. Grose and J. Ratner (editors), Human Factors and Web Development, Chapter 18, p241-254, LEA Press, 1997. Available at ResearchIndex (Citeseer): http://citeseer.nj.nec.com/245748.html

[Hascoet1999] Mountaz Hascoët: *Navigation and interaction within graphical bookmarks.* Rapport interne du LRI, NÂ°1232, 1999. (Internal Report of the LRI, No. 1232, 1999.) Available at ResearchIndex (Citeseer): http://citeseer.nj.nec.com/296052.html

[Hascoet] Mountaz Hascoët: *Integration of navigational aids in the user interface.* Hypertext'00, San Antonio, June 2000. Available at ResearchIndex (Citeseer): http://citeseer.nj.nec.com/323639.html

[Kaasten] Shaun Kaasten and Saul Greenberg: *Integrating Back, History and Bookmarks in Web Browsers.* In Extended Abstracts of the ACM Conference of Human Factors in Computing Systems (CHI'01), ACM Press, 2000. Accessible at: http://www.cpsc.ucalgary.ca/grouplab/papers/index.html

[Kahan] Jose Kahan, and Marja-Ritta Koivunen: *Annotea: an open {RDF} infrastructure for shared Web annotations.* In World Wide Web, 2001. pp. 623-632. Accessible at: http://citeseer.nj.nec.com/kahan01annotea.html

[Koch] Andreas Geyer-Schulz, Stefan Koch, and Georg Schneider: *Virtual Notes: Annotations on the WWW for Learning Environments.* Proceedings of the Fifth Americas Conference on Information Systems (AMCIS 1999), pp. 136-138, Milwaukee, WI, 1999. Available at ResearchIndex (Citeseer): http://citeseer.nj.nec.com/263780.html

[Laurent] Laurent, D., and L. Vignollet. *An annotation tool for web browsers and its applications to information retrieval.* In Proceedings of RIAO2000, Paris, April 2000. Accessible at http://www.univsavoie.fr/labos/syscom/Laurent.Denoue/riao2000.doc. Available at ResearchIndex (Citeseer): http://citeseer.nj.nec.com/denoue00annotation.html

[Li] Li, W-S., Q. Vu, E. Chang, D. Agrawal, Y. Hara, and H. Takano: *PowerBookmarks: A System for Personalizable Web Information Organization, Sharing, and Management.* In Proceedings of the Eighth International World-Wide Web Conference, May 1999. Available at ResearchIndex (Citeseer): http://citeseer.nj.nec.com/260729.html

[Lifantsev] Lifantsev, M.: *A System for Collaborative Web Resource Categorization and Ranking.* Ph.D. Dissertation Proposal, October 2001. Available at ResearchIndex (Citeseer): http://citeseer.nj.nec.com/496245.html

[Marais] Hannes Marais and Krishna Bharat: *Supporting cooperative and personal surfing with a desktop assistant.* In Proceedings of ACM UIST'97, 129--138, ACM, 1997. Available at ResearchIndex (Citeseer): http://citeseer.nj.nec.com/marais97supporting.html

[Maxwell] Maxwell, J. C.: *Beyond Cooperation.* 2002. Available at: http://www.byroncenterchurch.org/john_maxwell/beyond_cooperation.htm

[Pitkow] Pitkow, J.: *In Search of Reliable Usage Data on the WWW.* In Proceedings of the 6th International World Wide Web Conference, Santa Clara, CA 1997, pp. 451-462. Available at ResearchIndex (Citeseer): http://citeseer.nj.nec.com/242362.html

[Tauscher] Linda Tauscher and Saul Greenberg: *How people revisit web pages: empirical findings and implications for the design of history systems.* In Int. J. Human-Computer Studies 47, pp. 97-137, 1997. Accessible at http://ijhcs.open.ac.uk/tauscher/tauscher-pdf.html

[Terveen] Loren Terveen and Will Hill: *Beyond Recommender Systems: Helping People Help Each Other.* In HCI In The New Millenium, Jack Carroll, ed., Addison-Wesley, 2001. Available at ResearchIndex (Citeseer): http://citeseer.nj.nec.com/terveen01beyond.html

[Twidale] Michael B. Twidale, David M. Nichols and Chris D. Paice: *Browsing is a Collaborative Process.* Information Processing & Management, 33(6), 1997, 761-83. Available at ResearchIndex (Citeseer): http://citeseer.nj.nec.com/twidale96browsing.html

[W3CWCA] World Wide Web Consortium: *Web Characterization Terminology & Definitions Sheet.* Accessible at: http://www.w3.org/1999/05/WCA-terms/

[Wagner2002] Holger Wagner: *Towards an Approach to Collaborative Web Usage.* Accessible at: http://www.pms.informatik.uni-muenchen.de/publikationen/diplomarbeiten/Holger.Wagner/