Tagging Wikipedia: Collaboratively Creating a Category System

Katherine Thornton
The Information School
University of Washington
thornt@uw.edu

David W. McDonald
The Information School
University of Washington
dwmc@uw.edu

ABSTRACT
Category systems have traditionally been created by small committees of people who had authority over the system they were designing. With the rise of large-scale social media systems, category schemes are being created by groups with differing perspectives, values, and expectations for how categories will be used. Prior studies of social tagging and folksonomy focused on the application and evolution of the collective category scheme, but struggled to uncover some of the collective rationale undergirding the decision-making processes in those schemes. In this paper, we qualitatively analyze the early discussions among editors of Wikipedia about the design and creation of its category system. We highlight three themes that dominated the discussion: hierarchy, scope and navigation, and relate these themes to their more formal roots in the information science literature. We distill out four styles of collaboration with regard to category systems that apply broadly to social tagging and other folksonomies. We conclude the paper with implications for collaborative tools and category systems as applied to large-scale collaborative systems.

Categories and Subject Descriptors
H.5.3. Group and Organization Interfaces, H.5.4 Navigation

General Terms
Design, Human Factors.

Keywords
Categorization, Wikipedia, information organization.

1. INTRODUCTION
Trent was recently browsing Wikipedia and stumbled on a new page for the indie folk band Aquabats. Trent decided he would help by categorizing this page. He had never tagged in Wikipedia, but he had used del.icio.us before - how hard could this be? He looked for the category "Bands" and found it was redirected to another category. He found "Musical groups by genre" but couldn't find "Indie" or "Independent". He found "Folk" but couldn't categorize the Aquabats as "Folk rock" nor "Folk punk". He considered trying to add a new category, but a category with only one item wouldn't be all that useful. Thinking to himself that this was going to be harder than he initially thought, he gave up on the idea of helping.

GROUP'12, October 27–31, 2012, Sanibel Island, Florida, USA.

Social tagging became wildly popular with the advent of del.icio.us. Suddenly, the idea of visible and shared categories became popular with developers and researchers. Researchers could study how shared category metadata was applied, how it evolved, and how individual use differed from collective use. Yet, despite the large number of studies of tags and tagging behaviors, none have considered the collective rationale behind the tagging scheme because largely the rationale had to be divined from the tags themselves. This paper specifically considers the collective effort and negotiation that generated one tagging scheme currently in use and analyzes those negotiations.

Schemes for categorization or tagging have a direct impact on how people access information and how they can collaborate. These pieces of metadata can be used to help manage processes as tags are added and removed. Further, they form a type of communication system, a set of signals that collaborators can use. A better understanding of the rationale behind a tagging scheme will inform the future development of tools to improve the quality of the scheme, how the scheme is applied, and the way it provides access.

Wikipedia is an online encyclopedia created entirely of user-generated content. Roughly two years after Wikipedia began, the community decided to create a category system to organize and tag the content of the site. The category system has changed over time, as have conceptualizations of what role it should serve in Wikipedia. This study analyzes six months of discussion about the design and implementation of the category system in Wikipedia. The analysis reveals a set of important themes that related to category systems and tagging broadly.

In the following, we outline the prior work that considers tagging practices and relate that to some of the traditional work in information science. We describe our data collection, our approach to analysis and describe four key themes that emerge from the data. We then present a set of analytic frames for thinking about how category systems enable styles of collaboration and illustrate how individuals contributing to the discussion assumed these roles at different times. We close the paper with a short discussion and implications for collaborative systems.

2. TAGGING AND WIKIPEDIA
In Wikipedia, categories are applied to pages as an internal hyperlink to a 'category' page (see Figure 1). The WikiMedia system reads those special internal links and groups together all pages that include the link. Any number of category links can be applied to a page. The category system contains traditional content categories as well as management and historical categories that are often hidden when the page is requested. The simple application of a link to a page is similar to the metadata practice of adding a tag to a piece of content.

```
[[Category:American country rock groups]]
[[Category:American country music groups]]
[[Category:American folk rock groups]]
[[Category:Grammy Award winners]]
[[Category:Musical groups established in
1966]]
```

Figure 1. Example category links from the Wikipedia page "Nitty Gritty Dirt Band"

In an early analysis of tagging behavior, Golder and Huberman [5] argue that sensemaking is one purpose of tagging and that individual conflicts at the level of tag specificity makes collaborative tagging systems fuzzy and less precise. Subsequent studies by Sen et al. [12], Farooq et al. [2], and Millen et al. [10] served to illustrate the tension between individual tag choice and group tag use, quantitative techniques for considering whether a tag had information value, and the different motivations for tagging within an organization as compared to public tagging.

Our study extends what we know about tagging systems by characterizing the decisions about how a tag-based category system should operate, and how best to address these known issues through the design of the system.

The category system in Wikipedia has been considered from a number of different perspectives. Some studies follow a quantitative approach. For example, Muchnik et al. [11] tested five algorithms for automatic extraction of hierarchical relationships among Wikipedia pages against the extant category system. Kittur et al. [8] quantitatively mapped the categories used in Wikipedia into seven topical areas. This work collapsed the category system using a shortest-path-to-root approach for determining where node content would be assigned. This obscures the significant work Wikipedians have undertaken to systematize category hierarchy and encourage editors to apply category labels according to community convention. Holloway et al. [6] conducted a quantitative study of the category system from 2004 through 2007 to produce a thorough description of the category system as it stood in 2007.

As these studies show, quantitative approaches can illustrate the depth and breadth of the category system and compare the relative prevalence of topical content. But the quantitative approach cannot illustrate the collaborative rationale behind the system, and some quantitative approaches may specifically obscure the community efforts to apply category tags in a systematic manner.

Still, other studies have considered participation in Wikipedia from a qualitative perspective. Forte and Bruckman [3] outline social roles that exist in the Wikipedia community. They describe the negotiation that takes place in the wiki, and on related mailing lists between individuals who are trying to create community-wide policies. This highlights the negotiated aspects of much of the work in Wikipedia, including that around the category system. In a study of the types of work valued in the Wikipedia community, Kriplean et al. [7] point out that both category creation and category link application are types of work acknowledged and valued by the community. This work will support their argument by providing rationales for elaborating content organization advanced by the community in the early stages of the development of the category system. Analyzing discussions related to the creation and implementation of the category system, we describe strategies that the Wikipedia community employed in the attempt to make the implementation of category designations consistent throughout the site.

In the field of information science, knowledge organization (KO) theorizes, analyzes and critiques systems designed to organize information. From a KO perspective, Voss [15] described the Wikipedia category system as a thesaurus built through collaborative tagging. While Voss notes that simply adding categories to other categories facilitates the creation of hierarchy, he did not outline how different implied relationships are created, and sometimes confused, between categories when they are nested. Our analysis extends this finding by characterizing types of relationships present among super- and sub-categories in Wikipedia. There is surprisingly little evidence of the community prioritizing designs that would help users' distinguish between different types of relationships in the early discussions of the category system in Wikipedia.

3. METHODS

We collected the text of the Category talk page covering the six-month period from June 2004 through January 2005. That marked the start of the community effort to implement a category system in Wikipedia. At this point the design of the category system was completely open; changes were still being discussed and codified. We coded the data thematically in order to begin a content analysis of each section of the archived talk page. Through open coding we developed a codebook of 31 codes.

After coding the data and arranging codes into families, several themes emerged from the data. The editors who were participating in the discussions of how to design and implement the category system frequently returned to issues of hierarchy, scope, navigation and collaboration. By contrasting the editors' discussions of hierarchy with theory about constructing hierarchy in indexing languages, it is clear that many of the same conclusions reached through trial and error in Wikipedia have been described in the KO literature. The editors' discussions of how to scope the number of categories and the type of categories in the system are especially interesting when considered in the context of the recurring appeals to the community to consider extant schema for information organization as potential models. Discussions of navigation through Wikipedia via the category system reflect a very clear understanding of the category system as a navigational tool, something that no longer seems to be the case for many members of the Wikipedia community.

We then collected a second set of data (spanning January 2005-January 2007) from the same talk page and coded that data using the codebook. The primary difference between this data collection and coding process and the first round is that the number of codes relating to the purpose of the category system increased. Enough new codes relating to purpose were needed that we decided to introduce a new theme, that of 'purpose'. Our goal for collecting data from the second period was to learn how the community understood these themes as the category system was expanded and changed. We were also investigating whether new themes would emerge and how the focus of the community would shift over time. We will cover these themes in the following sections of the paper.

4. CATEGORY SYSTEM DESIGN THROUGH DISCUSSION

While working together to decide how a category system should function in Wikipedia, interested contributors shared their thoughts and debated the merit of many proposals. Several themes emerged from their discussion: hierarchy, scope, and navigation and collaboration. In terms of hierarchy, many editors felt that the best way to structure the category system would be through

category tags structured into hierarchical trees, and the WikiMedia code was designed to prompt users to always place a newly- created category into the supercategory of their choice. Scoping concerns are reflected in the way editors debated the appropriate number of category tags to apply to a given page in Wikipedia. Another theme addressed the question of whether navigation was considered as one of the purposes of the category system, and how best to address user needs related to navigation. Editors were also generally concerned with how the category system would facilitate forms of collaboration. We cover each of these themes in detail.

4.1 Hierarchy in the Wikipedia Category System

The community of editors who participated in the discussions was very concerned with ensuring that hierarchy would be a feature of the category system. The fact that the community felt hierarchy to be such an important element for organization is consistent with how the KO literature describes the advantages of hierarchical structures. Svenonius [14] noted that hierarchical relationships provide excellent advantages for supporting user navigation and the ability to instrument the relevance and size of a set of results for a user query. The fact that these advantages were apparent to the editors who designed the category system in Wikipedia is evident. When looking at the categories displayed at the bottom of any page it is possible to click on any of them (they are all hyperlinks) and see related categories and all subcategories. The editors who contributed to the design of the category system shared a vision of this feature that it would allow users to more quickly understand the context of any individual article by making relationships between articles visible.

Hierarchical structures allow users to see how different concepts relate to one another in a given system. Aitchison et al. [1] defines four types of hierarchical relationships: generic, whole-part, instance, and polyhierarchical.

Generic - A generic hierarchical relationship is defined as a conceptual transitive closure. There are very few examples of this in the category system. But there are other indexing tools in Wikipedia that are organized in this way, for example, the Wikipedia page for 'List of birds'. That is, if we take the class to be 'birds', all of the pages for which links are supplied in the list are pages for birds.

Whole-part - This relationship consists of a single concept or entity as the class with parts of that concept or entity as the subclass. An example of this type of hierarchical relationship in Wikipedia would be the pages listed under the category 'States of the United States'. Other than the page 'U.S. state', the other pages under this category are all parts of the category itself.

Instance - These are general concepts or classes which have specific instantiations as a subclass. It is difficult to find examples of categories for which the subcategories are all instances of the category. This is more often achieved through lists in Wikipedia. An example of this type of hierarchical relationship in Wikipedia would be the 'List of cathedrals' page. If we take 'cathedrals' to be the class, all of the cathedrals listed on that page are instances of the class.

Polyhierarchial - This type of relationship describes cases when one term is located underneath more than one category. Many categories in Wikipedia are located in more than one parent category. Polyhierarchy is very common in the category system of Wikipedia.

Much of the discussion in the data set illustrated a failure to discriminate between these different types of hierarchy, and multiple, sometimes contrasting, assumptions of what type of hierarchy was being proposed. All four types of hierarchy are in use in the category system of Wikipedia. In an ideal hierarchical structure a consistent relation would link terms. This recommendation is not followed in the category system of Wikipedia, and is likely a source of unaddressed, perhaps unrecognized, conflict in discussions of how the categories should be managed.

Browsing through the category system in order to examine the types of relationships that exist between superclasses and subclasses, the predominant relationship is one of association. Strictly, an Associative relationship is not hierarchical. The category system of Wikipedia, although perhaps envisioned to be a hierarchical system by some in 2004, is now full of categories related to other categories by an associative relationship. This is significant because although the designers felt that they were creating a hierarchical category system, many of the relationships in the category system are not hierarchical.

The fact that the relationships between supercategories and subcategories in Wikipedia include both hierarchical relationships as well as associative relationships is due to the fact that the category system emerged from a community in which there were divergent views of what the system should look like. One of the editors who contributed to the discussions in the dataset stated:

> So I think we need a way of distinguishing between a category where (a) you are asserting that everything in the category is an example of the thing it is in (ie list categories), and (b) categories where you are just providing hierarchical links for convenience. (editor 1)

The first type of category they describe encompasses the first three types of hierarchical relationships described by Aitchinson et al. [1], generic, whole-part and instance. The second type of category this editor describes would make use of the non-hierarchical, associative type of relationship. This editor is highlighting the need to be clear about the different types of relationships in the category system as it is being designed and created, and the differences this editor points to are elaborated in the KO literature.

Another editor provided the following example of why one page might need multiple category designations:

> I'm thinking about some of the dog topics. For example, dog is a member of pets; dog is also a member of mammals; both mammals and pets are members of animals but neither is a subcategory of the other. Now, how about dog agility? It needs to go under the dog sports category, which needs to be under the dog category, because it's related to dogs. It also needs to go under the sports category, because it's a sport. It probably also needs to go under the hobby category. But dog and sports do not at any higher point in the hierarchy have a common parent. (editor 2)

This statement is a clear articulation of the need for polyhierarchy. The editor would like there to be hierarchy, but would also like a single category to be able to belong to more than one superclass. This is an excellent example of the community working through

the issues until they come to a point of recognizing what they need. The advantages and limitations of this type of hierarchical relationship are well-documented in KO, and the Wikipedia community recognized the same issues when planning out the category system.

By 2006 the many editors were aware of the disjointed state of hierarchy in the category system:

> Welcome to the chaotic state of categorization on Wikipedia. As has been discussed extensively here and elsewhere, much of the chaos arises from two aspects: 1) there is no clear distinction between strictly hierarchical categories (article X is a type of/member of category Y) and associative categories (article X is somehow (often tangentially) related to category Y). The related to categories can turn into trails of free association. 2) Some editors will inappropriately remove an article from a parent category and instead place the article's eponymous subcategory within the parent category. This can lead to very strange hierarchies. Well, and aside from those two fundamental problems, there is always the problem of inexperienced users naively (mis-)applying categories. (editor 3)(period 2)

4.2 Scope of the Wikipedia Category System

Another theme that emerged from the discussions was the scope of the category system. The editors were very concerned about the number of category labels that would be applied to each page. One editor commented:

> Even if each of these categories is relevant (which can be doubted) the original page starts to clutter up rapidly. Logically, there is no almost limit to the extent to which categories can be applied to any page for the imaginative editor. (editor 4)

This editor was worried that categories would be assigned unevenly. Some individuals would choose to apply a large number of category labels, while others would apply few.

Some shared the concern that the number of categories applied to different articles would be widely divergent. One argument put forth was:

> I suggest that there should be a Guideline for categorisation by which editors (1) exercise caution and err on the side of not ascribing a category unless the text of the page justifies it (2)limit the size of the categorisation link text so that it remains small in relation to the size of the page. (editor 4)

While this editor clearly wanted to create hierarchical relationships between categories, they had observed how inconsistently these hierarchies were constructed.

Other editors felt that the work of scoping the category system of Wikipedia was such a large task that it should be modeled on existing structures for information organization. One editor brought up the challenge of making relationship types explicit and suggested modeling the category system on the Resource Description Framework (RDF).

> The fix is to label the arrows: describe the relations. This is, in my limited understanding, what RDF does. That uses the terms subject, predicate, and object. The subject is the thing you're categorizing. The object is the category you're adding it to. And the predicate describes the relation. Predicates allow you to make semantic inferences programmatically. (editor 5)

Other editors suggested modeling the category system on extant directories created to index the World Wide Web.

> To minimise reinvention of wheels, consider the category structures of Web directories such as www.zeal.com, which have been painstakingly thought out over long periods. (editor 6)

This editor is pointing out how much effort could be saved if the category system were modeled on an extant structure.

The design, creation, monitoring and evaluation of the category system required significant community effort. In order to create the most effective system, many discussions were based around how best to scope the category system. Members of the community expressed concern over inconsistency in the average number of categories that might get applied to a given page, made appeals to modeling the syntax of the system on RDF and suggested web directories as other potential models for the category system.

4.3 Navigation via the Category System

In the first period of data collection (June, 2004 through January, 2005) the hyperlinked category labels for each Wikipedia page were displayed at the top of each page. However, it is now the case that the category information is displayed in a box at the bottom of each page. There are many pathways to any individual page in Wikipedia. Users arrive to specific pages via a link from a results page from a search engine, from a link in the text of another page, from a link in an infobox located in the upper-left or upper-right- hand corners of many pages which typically contain pointers to a large amount of related content, from a list of links, or from the list of pages provided on the page for any category. At this time it is unclear how often the category system is used for navigation in Wikipedia. Regardless of the current reality, many of the editors who contributed to the design of the system in 2004 felt that navigation was a primary way the category system would be used.

One contributing editor articulated the following vision of navigating through the category system:

> We have to think from the encyclopedia user's point of view. He/she is starting at the top level of the hierarchy with a subject in mind, and they need to know which blind path to go down to find an article on that subject. It might help to think of the problem as a game of twenty questions. The first question we may ask is, "If they wanted to know about Stephen King's books, they might choose Category: Things, and have a choice of Category:Animals,

Category:Vegetables, Category:Minerals, Category:Ideas, etc., and go down one of those paths. My point is, Categories link only as a hierarchy; Wikipedia articles link as a network to every related article. So as long as the user reaches the article on Steven King (the person), or the articles on Steven King's books using the categories, the articles themselves link to each other. (editor 7)

This statement clearly indicates that the editor felt people would begin their search by looking at the category system from top to bottom for desired content.

Lee and Olson [9] compared the hierarchical navigation structure of Yahoo! directories with information retrieval via keyword searching in a search engine. One of the factors that they consider in their study is the location of the hierarchical browsing tool on the Yahoo! main webpage. They noted the harder it was to find the directory/ category information the less it was used.

From looking at the discussions in this dataset, it is clear that the decision to move the category box to the very bottom of each page predated community understanding of how the category system would be utilized.

Another editor presented a contrasting vision for how the category system would be used.

On the other hand, I think there's a good case to be made for a more bottom-up approach; let's take a look at how things are being categorized, and try to find the patterns in that. It's more the Wikipedia way, too. For example, I've noticed that there are a lot of categories that are non-plural, such as Category:Medicine, Category:Biology and Category:Law. In those cases, rather than being categories containing only one article (Medicine, Biology and Law, respectively) they are instead full of articles and subcategories that are about the indicated topic. (editor 8)

This editor is articulating a need for specific guidelines for term construction to facilitate vocabulary control, another example of the Wikipedia community echoing principles frequently discussed in the KO literature. This editor's comments suggest that some members of the community felt that the amount of effort that was being expended in the design of the category system could be reduced if the purpose of the category system were explicitly articulated.

4.4 Purpose of the Category System

In our initial period of data collection and coding we created several codes relating to the purpose of the category system. These codes were few in number and, in general, were used to label sections of text on the discussion page in which editors expressed desire to come to agreement on the purpose of the category system. In the second round of data collection, many editors expressed opinions about what they felt the purpose of the category system to be, thus we added more codes relating to purpose. This was such a frequently-discussed topic that we chose to add a fourth theme, that of purpose. All of the excerpts provided in this section are from the second period of data collection.

Some editors expressed that categories were tools for browsing:

Categories are intended to be an aid to browsing, rather than an general taxonomy (which would be POV and destined to fail). (editor 9)(period 2)

One editor stressed that their best use would be to support browsing, in particular, as opposed to search:

You don't need categories to find a particular article. Categories are not a search tool, they're a search-for-related-things tool. If you want to find a particular article, like Portland, Oregon, just type Portland, Oregon in the search box. (editor 10) (period 2)

Other editors raised the question of whether browsing was the primary purpose:

A question concerning the purpose of categories: Is the primary purpose of categories to: Aid the reader in finding material that may be of interest, or relevant to a particular topic? Producing a taxonomy; wherein being included in one or more categories is an indication—nay, a declaration by the Wikipedia community—that the subject of an article is an instance of the category it is included in. I seem to suspect the latter... (editor 11)(period 2)

Contrasting purposes were raised:

There seems to be a dichotomy between those who are looking to hone categories into encyclopedic taxonomies and those who are looking for a tagging system in which they can do keyword searches. The more we push at removing overcategorization, the more there is a need for a simpler tagging system. If we can answer that need, it might make everyone happier. (editor 12) (period 2)

The fact that multiple, sometimes conflicting, conceptualizations of the purpose of the category system were evident two years after the category system had been introduced is a challenge that the community frequently discussed.

5. MODES OF COLLABORATION AROUND CATEGORIES

Large-scale collaboration among editors of Wikipedia to create the category system is one of the primary differences between this system and development in the KO literature. Negotiations between editors take place around each decision that is made about the design and implementation of the system. While there are a far greater number of people contributing to the category system in Wikipedia than historically have labored over the conceptualization of a classification system such as the Dewey Decimal Classification (DDC), progress can sometimes be impeded by disagreements. In the discussions we observed several prominent themes related to collaboration. We identified four modes or styles of collaboration around the category system that were assumed by the participants:

- Collaboration with the category system - This theme describes discussion in which editors were conceptualizing the category system as an entity that facilitates navigation and or retrieval, clustering or conceptual visualization.

- Collaboration over the category system - This theme describes discussions in which editors were conceptualizing the category system as an object of work that individuals must use and manipulate. These include be debates about how categories should be applied, what types of relationships should exist and be made explicit between categories.

- Collaboration through the category system - This theme describes discussion in which the editors were conceptualizing the category system as a mechanism for communicating and interacting with others.

- Tools for category collaborations - This theme describes discussions in which editors discuss tools they are using to facilitate collaboration with, over and through the category system.

In the following we illustrate how participants in the discussions appealed to each of these styles or modes when making a case for features in the emergent category system.

5.1 Collaboration with the Category System

The theme of collaboration with the category system is applicable to discussions in which editors discussed how the category system will be engaged with by users. One contributing editor articulated the need to balance the workflow of editors who are applying categories with the needs of Wikipedia users.

People are creating categories from the bottom up because that's the easiest way for editors to work -- they put the four Beatles together in a category then lump them together into larger categories, because few people want to attempt to create a list of hundreds, or thousands, of articles. But however it's done, we have to make it easy for encyclopedia users to navigate from the top downward. Vegetarianism is fine within the discipline of Food and drink, as long as it isn't within a subcategory that's a list of foods or a list of drinks. (editor 7)

Another editor, replying directly to this comment, stated the importance of reducing the number of clicks users would be required to make in order to get to related content of interest.

I think we have a philosophical difference here. You rightly talk of the importance of top-down, and of a properly understood and maintained hierarchy. I agree completely. Where we disagree is that I think the wiki can encode much more than that (without breaking the behaviour you would like). It's clear that this is what people are trying to do, but only by breaking the hierarchy in the process. I also take the view that people are more likely to start in the middle of the hierarchy than at the top: they'll google their way to a Wikipedia article, spot the categories, and jump into the tree. Where do they go from there? From a user's perspective, categories are primarily a navigational tool. Frankly it would annoy me intensely if I surfed from a footballer to Category:Football (soccer) players and then had to start from the top of the hierarchy to reach Category:Football (soccer) rather than follow a single link. That link could go in a "Related links" section of the Football players category page, sure. But that's a kludge, I'm afraid, and throws away what could be meaningful data. I think the approach to take here is to add the ability to have relations in the category. It doesn't remove anything from the system that we have now and may add something. In the first implementation off this, all that should change is that the category page would have multiple lists, one for each relationship, rather than the single list it has now. This shouldn;t be too hard to code up and add extra possibilities without any downside (apart from the implementation time). (editor 13)

Another contributing editor highlighted the importance of categories as a support for browsing.

My boosterism of functional categories has been in support of that skimming, browsing user. Reading the above description, although intriguing, I must confess I have never considered the category tree as supporting that sort of precision data-mining search. Wikipedia strikes me as more of an imprecise, people-to-people exercise in information transfer, like any traditional encyclopedia in that sense. (editor 14)

Editors were very concerned with making decisions that would ensure the design of a category system to support navigation and retrieval, clustering of related content and conceptual visualization. In this way editors were designing a category system that would itself be a partner in the collaborative process of expanding Wikipedia.

5.2 Collaboration Over the Category System

The theme of collaborating over the category system is evident when editors discussed the category system as an object of work that individuals must use and manipulate. One editor expressed a need to find a solution to sorting issues within categories:

Consider the situation of a Category containing 30 articles with Sort Keys Book 01, Book 02, ..., Book 30. These would all appear under B using the current system; this would still hold if the threshold was measured against the number of articles as opposed to the number of "sort buckets". The latter is what I want to measure. The system which has been unilaterally adopted in Category:Harry Potter movies will only work for a series up to 9 items, since an article with Sort Key 10 will appear under 1 and screw up the sorting arrangement. I would prefer to use the system I originally installed, being more scalable, but am unwilling to impose it without some discussion. (editor 15)

This comment demonstrates concern for how design decisions will affect user experience and also serves as an appeal to other interested parties to help reach a majority opinion on how to mediate this issue.

Editors also discussed how the category system might help users conceptualize topics in relation to one another.

I'm not sure why "ease of maintenance" is an issue on this, or why that overcomes the great navigation and classification benefits that have previously been mentioned. Articles that define categories are not only the parents of those categories but will also logically be a member of whatever parent categories their own categories belong to. The articles should reflect this, for navigation purposes, as well as to properly classify the article. These are the two functions of categories. A reader of an article may not want to read just more topics on that article, but to see others of the same kind, and he may not even know that such a parent category exists without the article being tagged with it. Categorizing Ohio only under Category:Ohio just tells you that there are more articles about Ohio. A non-U.S. reader in particular may not assume that clicking through that may take him to other categories on other states, plus he may wonder why Ohio isn't classified as a U.S. state, if that's what the article tells him it is. Why unnecessarily increase the steps required to find what should logically be right there? Why omit classifications on the articles that are obviously the most (or all equally) important instances of that classification by virtue of their having a subcategory? (editor 16)

Concerns over how the category scheme as an object of work that individuals must use and manipulate required the editors who participated in the design and implementation of the category system to collaborate over the system design. This collaboration entailed providing scenarios with different outcomes in relation to issues of concern. It also involved explicitly soliciting the input of others before a decision would be implemented. We will discuss tools that were used to facilitate such calls for collaboration below.

5.3 Collaboration through the Category System

Collaboration through the category system is a theme that applies to discussions of how the categories themselves might serve as a mechanism for communicating and interacting with others. Feinberg [3] argued that knowledge organization tasks are vehicles for the expression of the creators' beliefs. The discussions about the Wikipedia category system expand her argument in the realm of systems in which the design has been massively collaborative. The participants must reconcile which points of view will be expressed through the organization and assignment of categories and how the resulting system affects adherence to the neutral-point-of-view policy.

Editors were concerned with how the assignation of categories might violate the neutral-point-of-view policy.

We as a group have begun in a few different places around WP to identify a potential problem with POV in categorization. It seems this is happening when a category is created that has a negative connotation and no self-evident criteria for inclusion/exclusion. (It probably could also happen with a category having an extreme positive connotation, but I haven't seen that come up yet.) I think it may be useful to understand more fully just which actual, current categories are subject to this phenomenon, so that we may draw better-grounded conclusions after inspecting a more full set of actual examples. To that end, I'm starting an alphabetical list here (feel free to chip in) of categories I think are likely to cause POV controversy. (editor 14)

This comment provides evidence that the community was aware of the potential for expression in the act of naming and applying a category. This discussion also contains a direct appeal for collaborative effort toward the end of identifying examples of categories that might be contentious.

Another example of how categories were seen to have the potential to be expressive of bias is evident in the following discussion:

I really don't like the idea of categorizing people by race, religion, or sexual orientation, so Category:Gay people should go, and Category:Jews and Judaism should be just Category:Judaism. Gay rights activists would be a proper category, however, as would Jewish religious leaders, as long as it is categorized by something someone does rather than what they supposedly are. I think categorizing people under Category:African Americans or Category:Asian Americans is highly offensive and POV. Whether someone is one or not is largely a matter of self-identification (how do you label yourself if you are multiracial?), and it is inherently POV to think that people are appropriately classified based on what race they are, as if that is a defining trait. It is much less offensive but still problematic to merely include this...information in list articles, because at least that way you're not slapping a classification on the subject of an article, saying "this is what he is". (editor 17)

5.4 Tools for Category Collaborations

As mentioned above, there were multiple points in these discussions where appeals for collaboration were explicitly made. Several tools were named that editors were using to facilitate this collaboration. On the topic of how category links are highlighted if they are created but left without being assigned to a parent category, one editor invited others to help address the problem.

If people created categories responsibly, there would never be a redlink category. A red linked category does not mean it doesn't have a parent. It means it has no description. (And if it has no description, it can't have a parent, but that's not the

point.) There are many criteria that determine if a category "exists"...does it have articles? Does it have a parent? Does it have an article (description)? Only the last of these means anything to the link color. If you hate red categories so much, maybe you'd like to join me in fixing them on Category:Orphaned categories where I have lots of them listed. (editor 18)

The function that such categories are highlighted with the color red could also be interpreted as a tool to support collaboration itself as it indicates to editors that the category lacks a parent, a vital part of the category assignation process. But as well, in the quote we see a range of ways to determine "existence" of a category that point toward a range of technical assistance for category creation and application.

Another way that editors collaborated was by sharing recommendations for how to accomplish certain types of tasks.

In the meantime, one useful method I've found is to go ahead and edit the article or subcategory you'd like to classify. Type in your best guess at the name of the proper category/ies, AND the name of a larger category that you're sure exists, which could be a (grand)parent (i.e. Category:Medicine or Category:Music, or Category:Musical groups by genre, as specific as you can get). Then use the SHOW PREVIEW, not the Save page button. Look for the previewed categories at the very bottom of the page (it may be below the Preview edit box, depending on your Preferences settings). If your best guess is blue, you've hit upon an existing category tree; if red, it doesn't exist or is spelled or worded differently. (editor 19)

This advice was a helpful work-around when the software did not support any type of browsing of the category system other than via an alphabetical list of all categories.

Editors also made direct reference to tools external to Wikipedia.

In Bugzilla, there is 'Bug 450: Categories need to be structured by namespace'. To some extent the lists are already structured by namespace (their name sorts them together). Contrary to the deletion log sometimes included in categories or user pages, images aren't just noise in the category, but informative. As it's easy for readers and other users to distinguish them from articles, I'd include them. As another example, one could quote Category:Saint Helena. (editor 20)

This comment indicates that some editors made use of the bug-tracking software, Bugzilla, to keep track of work that needed to be tackled within the Wikipedia project. Another editor mentioned Sourceforge:

RfE 964667 is probably the closest task in sourceforge and is currently unassigned. (editor 21)

This response was made to address a question of whether anyone had proposed creating a visualization tool that would display all dependent subcategories for a given category to facilitate the accurate assignation of categories to pages.

6. DISCUSSION AND IMPLICATIONS

Tagging is a valuable and popular collaborative technology. Many prior studies have focused on the application of tags without a clear tie to the underlying assumptions of the users who are applying those tags. The act of tagging, or labeling, an item as a member of some category of things or concepts has profound implications for the way we see or understand that item. In a collaborative context, the ability of the members to successfully negotiate disagreements over instances of labeling directly influences the progress and success of the collaboration.

The content of Wikipedia is connected as a graph structure with many pages explicitly linked to other pages using largely associative relations. Early on participants argued for the creation of a category system with clear hierarchical relationships. Many discussions of hierarchy did not distinguish between different types of hierarchical relationships, nor did they cover the challenges users would face when trying to interact with a system in which many different types of relationships would exist between categories without being explicitly described.

This points to specific opportunities for category tools. While Wikipedia has some extensions that allow for the exploration of the category system, there is nothing that specifically visualizes the assumed or real relationships among category nodes. A tool that in some way displayed the relation among categories would help regular users navigate with the category system and help individuals who wanted to tag pages.

Members of the community were also conscious of the issue of appropriately scoping the category system. They worried about consistency in terms of the number of categories and super-categories that might be applied to a page. They also suggested external models for the structure and syntax of the category system. They recognized that models and syntax structure would require work to create and maintain. Further, they also recognized that they might become a barrier to entry for newcomers who want to participate by categorizing uncategorized pages.

Scoping tools represent yet another possible technical enhancement. A scoping tool could help users understand whether a page might be over or under categorized. Similarity measures between sets of category tags and the text of pages could suggest new categories that might be applied or categories that may be unnecessary.

The category system was initially conceived as a navigational tool to complement traditional search. Many recommendations were made as to how to facilitate navigation through Wikipedia using the category system. It is unclear what role the placement of the category tags at the bottom of articles plays in the utilization of the system for navigation, but studies of other systems suggests that it may have a negative impact.

Wikipedia currently relies on search as the primary navigation system. But with all of the labor that has gone into the category system, an obvious enhancement would be to leverage the category system to provide users more contextual information about the content that they are viewing. Visual snippets of the categories relating to the page alongside their super- and sub-categories could help users understand a specific article or could

be attached to internal page links to present more context about the possible target page.

There is much future work to be done in this area. In particular, while there has been work to look at the relative distribution of tags (Kittur [8]) the growth of the category system itself has not been considered. Further how the growth of the category system mirrored or has not mirrored the application of category tags would also be important to know relative to existing studies of other collaborative tagging and folksonomic systems. The approach that Spinellis and Louridas [13] employed to examine the growth of Wikipedia via links to pages that do not yet exist and are subsequently created would be a useful model to explore for a similar study of the expansion of the category system.

7. CONCLUSION

Our study has analyzed the early efforts to collaboratively create a category system for Wikipedia. Through analysis of group discussions, we saw the collective concerns for how the category system would be structured, how it would be applied, and how it could be useful for future users of Wikipedia. Our analysis unpacks some of the collective and social concerns that individuals had about the creation and use of a category system.

Our analysis extends what we know about social tagging systems and the collaborative creation of category systems. The analysis of how the category system functions to enable different styles of collaboration is important in relation to the other collaborative tagging studies. This rubric of collaboration styles should be tested in other large-scale collaborative projects where category schemes are fabricated and used.

While we know from the work of Kriplean, et al. [7] that the labor invested in maintaining and improving the category system is valued work in the Wikipedia community, the system itself seems to be underutilized. If we harness the power of the information structures built by thousands of editors to provide context for each article among related content in order to present such relationships visually to users, this would provide an alternate navigational option with the potential to support sensemaking. If we are able to display category information that would support tagging decision making, we could encourage increased participation in the expansion and refinement of the category system, especially among novice editors.

8. ACKNOWLEDGMENTS

This material is based on work supported by the National Science Foundation under Grants IIS-0811210 and IIS-1162114. The opinions, findings, conclusions or recommendations expressed are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

9. REFERENCES

[1] Aitchison, J., Gilchrist, A., and Bawden, D. “Section F structure and relationships.” In *Thesaurus Construction and Use: A Practical Manual*. Chicago: Fitzroy Dearborn Publishers (2000): 49-84.

[2] Farooq, U., Kannampallil, T.G., Song, Y., Ganoe, C., Carroll, J.M. and Giels, C.L. Evaluating Tagging Behavior in Social Bookmarking Systems: Metrics and Design Heuristics. In *Proceedings of the ACM 2007 International Conference on Supporting Group Work*, (2007), 351-360.

[3] Feinberg, M. Expressive bibliography: personal collections in public space. *Knowledge Organization* **38** (2) (2011).

[4] Forte, A., and Bruckman, A. Scaling Consensus: Increasing Decentralization in Wikipedia Governance. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences*, (2008).

[5] Golder, S.A. and Huberman, B.A. The Structure of Collaborative Tagging Systems. *Journal of Information Science*, **32** (2). 198-208.

[6] Holloway, T., Bozicevic, M. and Börner, K. Analyzing and visualizing the semantic coverage of Wikipedia and its authors. *Complexity*, 12 (2007), 30–40.

[7] Kriplean, T., I. Beschastnikh, and D. W. McDonald. Articulations of wikiwork: uncovering valued work in wikipedia through barnstars. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, (2008), 47-56.

[8] Kittur, A., Chi, E. and B. Suh. What's in Wikipedia?: mapping topics and conflict using socially annotated category structure. In *Proceedings of the 27th international conference on Human factors in computing systems*, (2009), 1509-1512.

[9] Lee, H. L. and Olson, H. A. Hierarchical navigation: An exploration of Yahoo! directories. *Knowledge Organization*, **32** (1) (2005), 10-24.

[10] Millen, D.R., Feinberg, J. and Kerr, B., Dogear: Social Bookmarking in the Enterprise. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, (2006), 111-120.

[11] Muchnik, L., Itzhack, R., Solomon, S. and Louzoun, Yoram. Self-emergence of knowledge trees: Extraction of the Wikipedia hierarchies. *Physical Review E* **76**, 1 (2007), 016106.

[12] Sen, S., Lam, S.K.T., Rashid, A.M., Cosley, D., Frankowski, D., Osterhouse, J., Harper, F.M. and Riedl, J., tagging, communities, vocabulary, evolution. In *Proceedings of the 2006 ACM Conference on Computer Supported Cooperative Work*, (2006), 181- 190.

[13] Spinellis, D. and Louridas, P. The collaborative organization of knowledge. *Communications of the ACM* **51**(8), (August 2008), 68-73.

[14] Svenonius, Elaine. 2000. *The intellectual foundation of information organization.* MIT Press.

[15] Voss, J. Collaborative Thesaurus Tagging the Wikipedia Way, (2006) http://arxiv.org/abs/ cs/0-604036. Accessed Jan 21, 2011.