Joseph T. Tennis
The University of British Columbia, Vancouver, Canada

Comparative Functional Analysis of Boundary Infrastructures, Library Classification, and Social Tagging

Abstract: This paper outlines three information organization frameworks: library classification, social tagging, and boundary infrastructures. It then outlines functionality of these frameworks. The paper takes a neo-pragmatic approach. The paper finds that these frameworks are complementary, and by understanding the differences and similarities that obtain between them, researchers and developers can begin to craft a vocabulary of evaluation.

Résumé : Cet article présente trois cadres d'organisation de l'information : la classification des bibliothèques, l'étiquetage social et les infrastructures frontières. Cet article souligne les différentes fonctionnalités de ces trois cadres. Une approche néo-pragmatique est utilisée. Les résultats indiquent que ces cadres sont complémentaires et que par la compréhension des différences et des similarités qui existent entre eux, les chercheurs et les développeurs peuvent commencer à créer un vocabulaire pour l'évaluation.

1 Introduction

The purpose of this research is to identify distinguishing characteristics of three types of information organization frameworks: boundary infrastructures, library classification, and social tagging. Patrick Wilson outlines the function of library classification as identifying a *sense of position* (1968 p. 69-92), where a user can find all writings on a given subject in one place. In contrast, Bowker and Star propose a definition of classification as a *boundary infrastructure* in which they see "classification systems as historical and political artifacts very much as part of Western bureaucracy…including the conflicts of meaning that occur when multiple groups fight over the nature of a classification system and its categories," (Bowker and Star, 1999). Furthermore, thousands of Internet users currently employ what has been called *social tagging* or *folksonomies* (e.g., flickr, 2006). Information Science is just starting to examine the social tagging phenomenon, but in most cases information professionals see it as a weak and uneducated conception of information organization, identical to other traditional practices of subject cataloguing or indexing, but failing to control vocabulary and provide rigorous relationship structure.

Information organization frameworks possess similarities. Each works on a set of specifications and each creates categories in order to manage human-information interaction. However, each deploys these categories using different, and sometimes not explicit, specifications. This gives rise to differences in functions performed by these frameworks.

This paper examines the functions outlined by these three information organization frameworks – library classification, boundary infrastructures, and social tagging – in order to answer the question: are these three frameworks identical in purpose, form, and function? In order to address this question, a neo-pragmatic approach (Rorty, 1982) employing a reflexive hermeneutic analysis (Alvesson and Sköldberg, 2000) of texts and

instantiations of boundary infrastructures, library classifications, and social tagging was used. Three categories of functionality differentiating the three frameworks are reported: Lending Cognitive Authority (Wilson, 1983), Providing Transparency, and Fulfilling Tasks. Variations between these three categories of functionality make it possible to distinguish between library classification, social tagging, and boundary infrastructures.

It is concluded that these three information organization frameworks are complementary, without being redundant, and that information systems would benefit from employing multiple types of information organization frameworks – like social tagging along with library classification. As far as evaluation is concerned, this analysis could contribute in some small way to evaluating social tagging, boundary infrastructures and library classification for their intended and implemented functionality – while offering insight into improvements to such information organization frameworks.

1.1 Methodological Note

For the purposes of this study we approached the literature and instantiations of information organization frameworks (library classification, boundary infrastructures, and social tagging), from a neo-pragmatic perspective. Follow Rorty's discussion of the kind of knowledge social science should create; we followed his epistemological line of reasoning, and have made usefulness as the criteria for our statements (Rorty, 1982). We hope that the findings from this research will create a vocabulary that will help us cope with the social world, specifically the social world of information organization frameworks – their similarities, differences, and various criteria for usefulness. This approach has been acknowledged as a fruitful path in LIS (Sundin and Johannisson, 2005).

Within this neo-pragmatist perspective we applied a reflexive hermeneutic method (Alvesson and Sköldberg, 2000) to the literature and instantiations of the information organization frameworks. A reflexive hermeneutic method requires the researcher to examine the parts and the whole of these texts (including both the literature and the instantiations as texts), while acknowledging the researcher is part of a socio-political milieu. The literature consulted included the works cited in the reference list below.

This research is limited to a short discussion and only a few texts concerning these information organization frameworks. A more thorough analysis and discussion of these texts would allow us to make richer claims. Also, the criterion of utility, a claim from the neo-pragmatic epistemology taken by this paper, has yet to be invoked by anyone but the author. Now we will discuss the frameworks.

2 Information Organization Frameworks

We have argued elsewhere, that an information organization framework consists of information organization systems (classification schemes, taxonomies, ontologies, bibliographic descriptions, etc.), methods of conceiving of and creating the systems, and the work processes involved in maintaining these systems (Tennis, in press). Information organization frameworks surface when people need to manage human-information interaction. They are management tools. Retrieval is part of the management function of these frameworks, and it is a complex function. However, as we will discuss below, retrieval in these frameworks is not always the same function. In this paper, we will

compare three information organization frameworks here: library classification, social tagging, and boundary infrastructures. The next section defines these frameworks, and then outlines the components (systems, methods, and work processes) and functions.

2.1 Library Classification
Library classification is the practice and artifact of organizing written material so it is physically accessible based on subject (which may also include form or genre information, and the like). Traditionally, this was called bibliothecal classification to distinguish it from bibliographical classification (representations of the relationships between subjects in books without the need for physical and systematic shelf arrangement) (Richardson, 1912). S. R. Ranganathan provides a robust theoretical definition of library classification as an artifact. He says, library classification is:

      (a) the creation of a sequence that is helpful (by subject),
      (b) represented and fixed with ordinal numbers (drawn from a system of ordinal numbers),
      (c) designed to mechanize the maintenance of the sequence,
      (d) thereby creating a sequence of classes,
      (e) used in order to place any entity (known or as yet unknown) in a (helpful) position,
      (f) drawn from an infinite universe of entities, (Ranganathan, 1967, 77-78)

The practice of classification is a complex set of actions that includes the interpretation and then representation of the subject matter of an entity with an ordinal number.

The purpose of classification is to save time for the reader (user), by displaying these entities in helpful sequence – traditionally on the shelf – but more so now on the screen. The next section briefly introduces the systems, methods, and work processes involved in library classification.

2.1.1 Systems
In this case systems are classification schemes, for example the Library of Congress Classification (Library of Congress, 2005), Dewey Decimal Classification (Dewey at al., 2003), or the no longer updated, but still theoretically informative, Colon Classification (Ranganathan and Gopinath, 1987).

2.1.2 Methods and Work Process
The methods of Library Classification are two-fold: (1) creating the classification (2) using the classification to class entities (resources, documents, most often books). The former creates the artifact. The latter uses the artifact in aid to the process. Ranganathan created a method for doing both (Ranganathan, 1967). The practice of classification is different from the artifact. Wilson offers us insight into the practice of classification. To Wilson, classing an entity is to provide it with a *sense of position* (Wilson, 1968). He problematizes the interpretation process and the representation process, because, as he says, it is indeterminate. He argues that there is no way to determine *the* subject of an entity. Nevertheless, the framework for library classification requires a representation be made, after the interpretation of the entity in order

2.2 Social Tagging
A person engaged in social tagging labels web-accessible content using tags. Social tagging represents an extension of personal information management. It makes personal information management on the web public and, to a certain degree, shareable. The products of social tagging, folksonomies, are often collaboratively generated, open-ended labeling systems, that enable Internet users categorize content (Wikipedia contributors, 2006). Labels in these systems are called tags. Researchers speculate that the majority of the use of tags in the del.icio.us system is personal (Golder and Huberman, 2006, 207).

Folksonomies are collections of tags used in social tagging. They are folk (linked to people) taxonomies (nomenclatures used to identify categories). Folksonomies do not have any syndetic (thesaural relationship) structure. Researchers have identified seven types of tags created in folksonomies on del.icio.us. These folksonomies contain tags that:
> (1) identify what or who the resource is about
> (2) identify what the resource is, its genre (e.g., article, blog, or book)
> (3) identify who owns the resource
> (4) refine categories (often using numbers, e.g., 25, 100)
> (5) identify qualities or characteristics of resources (e.g., scary, funny, stupid)
> (6) reference to one's self (e.g., mycomments, mystuff)
> (7) organize tasks (e.g., toread, jobsearch) (Golder and Huberman, 2006, 203).

Folksonomies are identified as a collection because they are linked to a profile on the web. It is the profile that sets social tagging apart from boundary infrastructures and library classification. Profiles may be robust or thin in their description of the person responsible for it. For example Connotea (Connotea, 2006a), a social tagging system for scholars, provides profile space for its users. Here users can provide as much or as little information as they wish. Some users have used this feature to provide their full names, occupation (e.g., PhD student), research interests, affiliation (e.g., University of Cambridge), contact information, and other personal information like stating: *I would never leave my house without my…,* and *about connotea.* The profile also points to the user's Connotea collection of links, and in some cases to blogs, skype (over-internet telephone) account name, and messaging account name.

At the time of writing there are 51 profiles on Connotea. Many of them are thin on information – pointing only to the collection of links affiliated with this profile. Others are robust, containing the different types of information above and more.

Folksonomies, as collections of tags, act in a similar fashion to post-coordinate descriptors with varying degrees of exhaustivity and specificity. The variation is due to the very personal nature of interacting with this framework.

2.2.1 Systems
Folksonomies function as the system used in social tagging. Folksonomies are to a degree collaborative, but they are not completely collaborative. They also have come under criticism from some that without any kind of structure, these systems cannot support effective retrieval. As a result some have posited the need for *collabularies,*

which combine folksonomic systems with controlled vocabulary systems (Lawrence, 2006).

2.2.2 Methods and Work Process
The methods employed by current social tagging frameworks are personal. Users of these frameworks exhibit different frequency of use, numbers of tags created, and kinds of tags created (Golder and Huberman, 2006). Though the folksonomy systems used in social tagging frameworks are called collaborative, the decision to use other peoples' tags is a personal one, based not on any standards or shared understanding, but rather on a *folk method* of tagging, categorization, and personal information management. It is possible that further study will find significant differences between the methods of personal information management (Bruce, 2005) and social tagging, but at first blush they look very similar.

User requests drive methodological innovation in systems that support social tagging. As users of social tagging systems (e.g., del.icio.us, Connotea, flickr) request changes to functionality, systems designers respond. Connotea has a wiki for requested features (Connotea, 2006b), and del.icio.us has a discussion list (del.icio.us, 2006b).

2.3 Boundary Infrastructures
Boundary infrastructures are information organization frameworks that connect boundary objects. They knit together relationships between people, things, moral order, categories and standards (Bowker and Star, 2000, 286). Boundary infrastructures exist as coordinated work, but work that is from different communities. The Nursing Intervention Classification is an example of a boundary infrastructure. As a boundary infrastructure it connects medicine and its conceptions of work and knowledge with nursing and its conceptions of work and knowledge. Bowker and Star…

Bowker and Star tell us that boundary infrastructures are often hard to see, and require a particular sensitivity to power dynamics in the classification of objects and information in information systems.

2.3.1 System
The systems deployed in boundary infrastructures are many and diverse. A boundary infrastructure coordinates these types systems: classifications, multiple communities, work tasks, and often-heterogeneous sources of information.

2.3.2 Method and Work Process
Boundary infrastructures do not use formalized methods. This framework, like social tagging, is informal. Boundary infrastructures accrete methods, combined with work practices, in a complex overlay of work, information, and power positions (Bowker and Star, 2000).

Work processes involved in the development and use of boundary infrastructures are not fully explored in Bowker and Star (2000). Reviewers have found this a key point lacking in their discussion (Fidel, 2000). Bowker and Star point to several areas of future research, and this is one of them.

3 Functionalities: Lending Cognitive Authority, Providing Transparency, and Fulfilling Tasks

These three information organization frameworks possess commonalities and differences. In common, they share the creation and use of categories to manage human-information interaction. All three also work from a set of specifications, and though the exact nature of these specifications may be different between them, it is the presence of specifications that allows us to see how they are similar and different.

Because they are built on specifications, these frameworks can be analyzed using a functional point of view. A functional point of view asks: what functions do these frameworks enable? Are these functions designed or created by accident? We will address three functions present at different levels in these frameworks: Lending Cognitive Authority, Providing Transparency, and Fulfilling Tasks.

3.1 Lending Cognitive Authority

Cognitive authorities are the people who know what they are talking about, when they talk about something beyond our own experience (Wilson, 1983, 13). Wilson provides us with a rich description of cognitive authority (Wilson, 1983). In so doing, he separates authority from expertise. For example, an astrologer may be an expert in the doctrine in astrology, but we may not count them as an authority (Wilson, 1983, 27). Cognitive authority is a relationship between at least two people, and it is not enforced or enforceable. The basis for cognitive authority comes in a variety of ways. They can be educated so that they become experts, or possess the ability to find something out. Some may have a reputation as being experts and by extension authorities on a matter.

Information organization frameworks can display (or make transparent, see below) various components that may add or decrease authority in the management of human-information interaction. Social tagging, with its profile component provides information so that individuals can make assertions about whether or not to treat a tagger as a cognitive authority. The faceless and nameless library classification does not lend any additional features of cognitive authority to the systematic arrangement of classes in a helpful sequence – the entity you are looking for is either there or it is not. Lending cognitive authority is a function not built into this framework. Boundary infrastructures, because they are sites of (sometimes creative) tension between different communities seems not to work toward any positive use of cognitive authority. In the case of the NIC, nurses' knowledge works against the established and commonly held cognitive authorities – the body of medical knowledge.

3.2 Providing Transparency

Transparency is the ability to see the information organization framework and its systems, methods, and work practices. In library science, researchers and practitioners argue over whether information organization frameworks should be transparent – whether users should see the work practice or rules of library classification for example. Because librarians are engaged in a delegated activity – an activity done on behalf of someone else – more is hidden. For example, does it matter to a user that the Vancouver Public Library works from three different versions of the Dewey Decimal Classification? However it matters that there are three different sets of schedules to work from when one is subject

cataloguing in this organization. And Sauperl has found that some cataloguers are engaged in creating and representing meaning for themselves, and favouring it, even if they are aware of a user's meaning which may be different (Sauperl, 2004).

Boundary infrastructures suffer from opposing priorities of the function of providing transparency. Local actors involved with the boundary infrastructure may see what others consider invisible – but they do so because power relations affect their actions in when interacting with this framework. Social tagging is a completely transparent affair. Because the methods and work practices are almost entirely personal

Social tagging is very transparent. Users engaged in this act of creating categories are themselves establishing specifications. It is not a case of delegation in social tagging. Nor is there a concern, in the personal use of tags, for voice, the way boundary infrastructures struggle with making voices heard in a cacophony of conflicting categorizations, each playing a part in a potentially socio-political tension.

Transparency is linked directly to specifications. If an information organization framework has clear specifications, and if these are exposed to all who are engaged in the human-information interaction, then there is transparency.

3.3 Fulfilling Tasks

Tasks function at the intersection of categories and specifications in information organization frameworks. A traditional task for library classification is collocation by subject. Ranganathan would add – collocation by helpful sequence. Collocation is an example of a unitary class, but there are other types of tasks in information organization frameworks, including work-arounds and localization tasks. We will consider unitary tasks first.

Unitary tasks are feats of massive engineering. The complex interaction between information and users, if it is to be of quality, requires rigorous and systematic design and implementation. Ranganathan's *Prolegomena* is a 640-page introduction to the techniques of creating a fully faceted classification – for arranging subjects in a helpful sequence through a system of ordinal numbers. As Bowker and Star note – these kinds of knowledge-intensive and work-intensive engineering practices admit, in practice, to *residual categories* where principles are not uniformly applied (Bowker and Star, 2000, 10-11). That is, the real world of these information organization frameworks - especially in designing for fulfilling tasks – is messy. Therefore, careful thought must go into accounting for residual categories in order fulfill a unitary task consistently.

Boundary infrastructures exhibit a complex set of tasks – many of which are local against the standardization of massive systems. Actors interacting with these frameworks may work around the established tasks to subvert the purpose of the classification. For example, nurses want to both establish categories of direct care (administering care to a patient) and indirect care (checking charts during each round). This reconfigures the task of early designs of the NIC, because its designers wanted it to primarily be about direct care (Bowker and Star, 2000, 245-246).

Social tagging, as an extension of personal information management, is a complex set of tasks that are locally defined. Researchers have found social tagging to be a process of sense-making (Golder nad Huberman, 2006; Weick, Sutclifee, and Obstefeld, 1991). In this light, the primary task is making sense of what is found. It is also possible that the intended task is related to *personal, anticipated information need* (Bruce, 2005). This would admit to a diversity of tasks proposed by the tagger.

Fulfilling tasks then is function of information organization frameworks that can be accidental and based on reconfigured work practices, or it can be engineered to create consistency and to eliminate residual categories. Fulfilling tasks can be local and personal in nature.

4 Identity and Divergence in Information Organization Frameworks
Boundary infrastructures, library classification, and social tagging are all frameworks for managing human-information interaction. They all create categories based on specifications (explicit or otherwise). However, beyond this general purpose, they are quite diverse in form, function, and purpose. They each handle cognitive authority, transparency, and task fulfillment differently.

5 Conclusion
From the above discussion, we conclude that library classification, boundary infrastructures, and social tagging are diverse sets of category systems. Though some would explain all of these as weaker or stronger retrieval devices, we have simply tried to lay bare the device. That is, to describe these frameworks, their functions, and components, and in so doing highlight their diversity, yet retaining their commonality. The utility in doing this is to begin to craft a new vocabulary, which will allow us to begin discussions of implementation and evaluation.

A recent report from the University of California Libraries has raised questions relating to library classification and social tagging (BSTF, 2005). Likewise, a researcher at Cornell has called for a change in the library catalogue, which has engendered a rebuttal (Mann, 2006). The crux of these reports (and the rebuttal), lies in an understanding of what information organization frameworks, like library classification, are constructed and what their purpose is. The three frameworks outlined in the paper are complementary. They are not built for the same purpose, and they do not exploit identical functionality – they are not redundant pieces of work. If multiple types of functionality were desirable to implementers, like the University of California system, or even the Library and Archives Canada in their Metadata Catalytic Initiative, then frameworks could be combined.

The final benefit from conceptualizing information organization frameworks as having diverse functionalities, and explicating them here, is to begin the work of evaluation, where we ask *how well does social tagging work?* With a vocabulary in place for discussing the components and functions of frameworks like social tagging, then we can dissect that question, and provide clear research questions that can begin to address them.

References

Alvesson, M. and Sköldberg, K. 2000. *Reflexive Methodology: New Vistas for Qualitative Research*. (London: Sage Publications)

Bowker G. and Star, S. L. 1999. *Sorting Things Out: Classification and it Consequences.* Cambridge, Massachusetts: MIT Press.

Bruce, H. 2005. Personal, Anticipated Information Need. In *Information Research* 10(3). Available: http://informationr.net/ir/10-3/paper232.html

BSTF [Bibliographic Services Task Force]. 2005. *Rethinking How We Provide Bibliographic Services for the University of California.* Available: libraries.universityofcalifornia.edu/ sopag/BSTF/Final.pdf

Calhoun, K. 2006. *The Changing Nature of the Catalog and its Integration with Other Discovery Tools.* Available: http://www.loc.gov/catdir/calhoun-report-final.pdf

Connotea. 2006a. Available: http://www.connotea.org/

Connotea. 2006b. RequestedFeatures. Available: http://www.connotea.org/wiki/User:RequestedFeatures

del.icio.us. 2006a. del.icio.us. Available: http://del.icio.us

del.icio.us. 2006b. Feedback. Available: http://del.icio.us/help/feedback

Dewey, M., Mitchell, J. S., Beall, J., Martin, G., Matthews, W. E., New, G. R. 2003. *Dewey Decimal Classification and Relative Index.* (Dublin, OH: OCLC).

Fidel, R. (2000). Review of *Sorting Things Out: Classification and Its Consequences* by Geoffrey C. Bowker and Susan Leigh Star. In *Journal of Academic Librarianship* 26(5): 370-371.

flickr. 2006. Available: http://www.flickr.com/

Lawrence, A. 2006. The Evolution of Search. *Information Age*. Available: http://www.information-age.com/article/2006/march/the_evolution_of_search

Library of Congress. 2005. Library of Congress Classification Outline. Available: http://www.loc.gov/catdir/cpso/lcco/lcco.html

Mann, T. 2006. *The Changing Nature of the Catalog and Its Integration with Other Discovery Tools. Final Report. March 17, 2006. Prepared for the Library of Congress by Karen Calhoun.* Available: http://www.guild2910.org/AFSCMECalhounReviewREV.pdf

Ranganathan, S. R. 1967. *Prolegomena to Library Classification.* 3rd.

Ranganathan, S. R. and Gopinath, M. A. 1987. *Colon Classification* 7th edition. Bangalore: Sarada Ranganathan Endowment for Library Science.

Richardson, E. C. 1912. *Classification, theoretical and practical, together with an appendix containing and essay towards a bibliographical history of systems of classification.* New York: Scribner and Sons.

Rorty, R. 1982. *Consequences of Pragmatism*. Minneapolis: University of Minnesota Press.

Sauperl, A. 2004. Catalogers' Common Ground and Shared Knowledge. *Journal of the American Soceity for Information Science and Technology* 55(1): 55-63.

Sundin, O. and Johannisson, J. 2005. Pragmatism, neo-pragmatism and sociocultural theory: Communicative participation as a perspective in LIS. *Journal of Documentation* 61(1): 23-44.

Tennis, J. T. In press. Function, Purpose, Predication, and Context of Information Organization Frameworks. *Knowledge Organization for a Global Learning Society: Proceedings of the 9th International Conference for Knowledge Organization*, ed. Budin, Gerhard. International Society for Knowledge Organization 9th International Conference. Vienna, Austria. Jul, 2006. Ergon. Würzburg.

Weick, K., Sutcliffe, K. and Obstfeld, D. 2005. Organizing and the Process of Sensemaking. *Organizational Science* 16(4): 409-421.

Wikipedia contributors. 2006. Folksonomy. In *Wikipedia.* Available: http://en.wikipedia.org/w/index.php?title=Folksonomy&oldid=49234698

Wilson, P. 1968. *Two Kinds of Power: Essays on Bibliographical Control.* Berkeley, CA: University of California Press.

Wilson, P. 1983. *Second-Hand Knowledge: An Inquiry into Cognitive Authority.* Westport, CT: Greenwood Press.