Local Authorial Voice and Global Authorial Voice in Community-Authored Knowledge Organization Systems

Chris Holstrom
University of Washington iSchool
cholstro@uw.edu

ABSTRACT

Folksonomies are crowdsourced knowledge organization systems that rose to popularity during Web 2.0 and that are still actively used today. This crowdsourced approach to knowledge organization moves authorial voice from an individual expert or small group of experts to the community. What does it mean to have many voices contribute to a knowledge organization system? Do community members create a collective authorial voice? Are minority opinions more readily included? How does access to information, especially "long tail" information, change? This paper explores these questions by examining authorial voice in community-authored knowledge organization systems (CAKOS) and expert-authored knowledge organization systems (EAKOS).

Keywords

Authorial voice, knowledge organization system, folksonomies, classification.

INTRODUCTION

Descriptions of folksonomies focus on their low barrier to entry and inclusiveness, noting that the "bottom-up" design of folksonomies runs inherently counter to EAKOS (Vander Wal, 2007). However, inclusive design does not necessarily lead to traditionally marginalized voices gaining authorial voice. David Sturtz (2004) notes that the implications of classification remain problematic for minority voices, even in folksonomies:

> The democratic approach of a folksonomy avoids many of the ethical and political concerns of top-down, centrally-imposed systems. It allows the users of the system to establish their own sense of balance within the system, to use their own vernacular for indexing and retrieval, and prevents exclusion by creating new categories as needed. However, with no guiding hand at the helm, this communal approach has the power to shut out unpopular or misunderstood terms. The will of the community may flood them with useless content, use them in an unintended way, or marginalize them so that they essentially disappear.

Feinberg (2006) furthers Sturtz's concerns, questioning whether folksonomies are really democratic. She cites a lack of synthesis and community cooperation in CAKOS: "This characterization does not extend to any sense of the community coming together to determine how resources should be indexed, even by voting…social classification is more like libertarianism, where everyone's whims are allowed to flourish."

While Feinberg and Sturtz make clear that even a traditionally marginalized voices *can* be heard in CAKOS, they differ on whether these long-tail voices "flourish" or "essentially disappear." The difference between these seemingly polar viewpoints could be ascribed to differing views on authorial voice.

LOCAL AUTHORIAL VOICE

Feinberg's definition of authorial voice in knowledge organization systems stems from literary authorial voice (2011), which privileges expert authors with limited scope and aligns with what I call *local authorial voice*. Local authorial voice in CAKOS has a limited scope, a limited audience, or a limited number of authors. Local authorial voice is well-suited to representing minority or marginalized viewpoints. For example, a small group of ukulele hobbyists might use their own vocabulary to tag fingerpicking resources. The broader community might never see these tags or resources--even the local use of popular tags like "tutorial" might barely register globally--but these tags and resources are highly useful within the locale of the ukulele community and are still present in the broader folksonomy.

Many local authorial voices can exist independently and simultaneously in CAKOS. For example, Coqc (2015) considers the effectiveness of folksonomies in giving voice to endangered languages, studying a small group of Twitter users who tweeted in Sami and used Sami hashtags. Coqc finds that "Twitter has indeed the potential for being a valuable instrument for small localized groups and endangered languages. It can increase language use and contribute to the visibility of languages that are not present in mainstream and traditional media." Notably, Coqc cites the value to "small localized groups" and does not claim that Sami hashtags have a major presence on Twitter as a whole.

What we see, then, is that CAKOS support local authorial voice by allowing underrepresented voices to more readily *exist* in classification schemes than do EAKOS. But they do not necessarily make minority voices significantly more *visible*. This lack of visibility--often manifest as a lack of presence in navigational elements or in difficult-to-discover tags-

-can hinder access to minority viewpoints beyond specific locales.

GLOBAL AUTHORIAL VOICE

Visibility and ease of access come with *global authorial voice*. Numerous studies (Kipp & Campbell, 2006, Munk & Mork, 2007) have shown that the distribution of tags in folksonomies is Zipfian. The most popular tags represent an overwhelming majority of the tags and, arguably, an overwhelming majority of authorial voice. These heavily used tags, in aggregate, form what I call global authorial voice. Global authorial voice in CAKOS has broad scope spanning a large collection or set of collections, a broad audience, and many contributors. For example, all de.licio.us users' "music" tags represented a global authorial voice.

Global authorial voice in CAKOS is much more problematic for minority voices than local authorial voice. Unlike local authorial voice, global authorial voice represents an aggregation or, as Feinberg (2006) astutely observes, an unsynthesized mass, of many voices. Most folksonomies take majority-rule approaches to global authorial voice, favoring the most popular tags in navigation and autocomplete features. The realities of global authorial voice in CAKOS leave us with two questions: 1) Can a community produce a coherent authorial voice through consensus or some other means? and 2) Can CAKOS give meaningful visibility to minority voices?

Global Authorial Voice and Coherence

Coherent knowledge organization schemes are historically attributed to a single authorial voice and purpose. As Broadfield (1946) notes, "Classification only becomes possible when I decide what to do with it, and a single definite purpose emerges." CAKOS rarely have a single purpose and by definition do not have a single author. Halpin et al. (2007) agree with Feinberg's assessment that coherence cannot emerge without synthesis, stating, "because users are not under a centralized controlling vocabulary, no coherent categorization scheme can emerge at all from collaborative tagging."

However, CAKOS have evolved. Some of the most successful and coherent CAKOS embrace social tagging, but rely on moderators and editors--often members of the community themselves--to shape contributions into a coherent whole. For example, Stack Overflow allows anyone to tag their own question, but still maintains a coherent and navigable classification scheme through synonym merging, community moderation, and editorial standards for tags. The Stack Overflow scheme is not all that different from LCSH in process or coherence. Neither has a single authorial voice, but they achieve some level of global coherence by communicating and enforcing standards. What this suggests is that CAKOS that embrace freewheeling folksonomic ideals fail to achieve coherence because they lack synthesis, but that modern CAKOS can achieve coherence by adding layers of editorial control and community engagement to their tagging and maintenance processes.

Global Authorial Voice and Minority Voices

Although minority voices, such as Sami language users, can establish a local authorial voice, establishing a visible and global authorial voice is much more difficult. In practice, member dynamics in CAKOS, particularly: 1) who participates and how extensively, 2) how the broader community is managed, and 3) how the CAKOS is designed to surface minority voices, have a significant effect.

Bates and Rowley (2010), in comparing UK public library models for inclusion with LibraryThing's folksonomy, find that LibraryThing "offers benefits over LCSH... in the discoverability and representation of LGBTQ resources. However, the folksonomy is dominated by US taggers, and this impacts on the tagging of ethnic minority resources." So, while CAKOS can provide global authorial voice to some historically marginalized groups, the demographics and dynamics of the tagging community might not lead to greater representation for other groups. Simply allowing anyone to tag resources is insufficient to represent marginalized voices; you also must actively build communities that foster their contributions and design platforms that make those contributions visible.

While it is generally admirable to build a CAKOS to give more global visibility to minority voices than an EAKOS can offer, it is important to note that promoting minority voices is not always a good thing. Fringe hate groups and nefarious actors have intentionally over-represented their voice in the digital world by using bots and taking advantage of the platform that relatively open social and crowdsourced forums provide. Robust community management and thoughtful editorial control are even more important in this age to ensure that organic minority voices can bubble up instead of fraudulent, harmful, and spammy voices looking to gain the influence that global authorial voice grants.

CONCLUSION

While most CAKOS allow open participation, they do not necessarily allow minority voices to achieve greater visibility in classification schemes than do EAKOS. Minority voices can readily exist in CAKOS, but are most often limited to a local authorial voice. CAKOS that are more successful at promoting and integrating minority voices rely on editorial control and moderation to build community. Strong community can provide the synthesis necessary to promote minority voices and develop coherent schemes, delivering on the promise of democracy that folksonomies first offered.

REFERENCES

Bates, J., & Rowley, J. (2011). Social reproduction and exclusion in subject indexing: A comparison of public library OPACs and LibraryThing folksonomy. *Journal of Documentation 67*(3), 431-448.

Broadfield, A. (1946). *The philosophy of classification*. London: Grafton.

Cocq, C. (2015). Indigenous voices on the web: Folksonomies and endangered languages. *The Journal of American Folklore 128*(509), 273-285.

Feinberg, M. (2006). An examination of authority in social classification systems. *Advances in Classification Research Online 17*(1), 1-11.

Feinberg, M. (2011). How information systems communicate as documents: the concept of authorial voice. *Journal of Documentation 67*(6), 1015-1037.

Halpin, H., Robu, V., & Shepherd, H. (2007, May). The complex dynamics of collaborative tagging. In *Proceedings of the 16th international conference on World Wide Web* (pp. 211-220). ACM.

Kipp, M. E., & Campbell, D. G. (2006). Patterns and inconsistencies in collaborative tagging systems: An examination of tagging practices. *Proceedings of the American Society for Information Science and Technology 43*(1), 1-18.

Munk, T. B., & Mork, K. (2007). Folksonomy, the power law & the significance of the least effort. *Knowledge organization 34*(1), 16-33.

Sturtz, D. N. (2004). Communal categorization: the folksonomy.