# Finding social roles in Wikipedia[1]

Howard T. Welser, Dan Cosley, Gueorgi Kossinets,
Austin Lin, Fedor Dokshin, Geri Gay, and Marc Smith

**Abstract:**

The concept of "social role" has long been used in social science describe the intersection of behavioral, meaningful, and structural attributes that emerge regularly in particular settings and institutions. We use structural signature methods to identify key roles in a large distributed collaboration system (Wikipedia) by examining the distribution of edits across types of pages and the structure of relationships between editors. We distinguish between technical editors, substantive experts, vandal fighters, and social networkers and demonstrate key ways that their patterns of interaction and contribution differ. We conclude by considering how differential entry into and retention in particular roles may affect the operation of the large social system.

## 1. Introduction

It is difficult to gauge the social significance of Wikipedia, other than to say that, like Google, Wikipedia has forever changed how we use, find and think about information. Both pundits (like Stephen Colbert) and researchers (Stvilia et al. 2005, Giles 2005) have been pre-occupied with the question of whether the product of Wikipedia is of sufficient quality, or whether its pages constitute legitimate references (Read 2006). Others argue that the project will never achieve a sufficient level of quality relying on non-expert volunteers of unknown identity (Chesney). Instead of prognosticating about the potential of the Wikipedia project we contend that scholars should be more focused on understanding how Wikipedia has achieved the success that it has: a "pretty good" resource for a basic understanding of most any topic. How has the "pretty good" and incredibly extensive resource been achieved? And how has this been possible given the absence of the resources and controls of conventional firms and bureaucracies?

Through the course of our research we have come to suspect that the success of Wikipedia has stemmed from two key sources: (1) infrastructural features that help people find their roles in the organization (2) technical innovations that allow substantial economies of scale in the performance of many of those roles. This paper concentrates on the former and is about finding roles in a double sense: to what extent can we identify the roles people play in Wikipedia by measuring general behavioral and structural features of their participation? We address this technical challenge by using a range of qualitative and quantitative methods to find signatures of social roles in Wikipedia (cite cite). The second sense asks: how do people find *their roles* in Wikipedia? We address this question quantitatively by comparing general patterns in editing across two different samples of Wikipedians, a cohort that first edited in January of 2005 and a sample of dedicated Wikipedians whose edits spanned a period of greater than one year. These samples allow us to compare how patterns in participation vary between cross section of new participants and those who stick around, and become the long term participants in the social system.

## 2. Finding Social Roles

### 2.1 Roles in social interaction, online, and in Wikipedia

Across social settings we can identify people who are playing social roles: advisors, parents, brokers,

---

[1] Direct all correspondence to Howard T. Welser, welser@ohio.edu; htwiii@gmail.com. Part of this research was supported by the National Science Foundation under Grant No. SES-0537606.

editors, managers, or vandals. The concept of "social role" has long been used in social science describe the intersection of behavioral, meaningful, and structural attributes that emerge regularly in particular settings and institutions (Parsons 1951; Merton 1968; Callero 1994).

Social roles have mainly been studied online in the context of text based discussion spaces, where a variety of roles have been identified: local experts, answer people, conversationalists, fans, discussion artists, flame warriors, and trolls (Burkhalter and Smith 2003; Golder 2003; Turner et al 2005; Herring 2004; Haythornthwaite and Hager 2005). Insight into these social roles has been gained through ethnographic study of the content of interaction (Golder 2003; Donath 1996; Marcoccia 2004) and through the use of behavioral and structural cues (Viegas and Smith 2004; Turner et al 2005).

Wikipedia differs from discussion spaces in that the primary activity of the community is the construction of an artifact. In this way, Wikipedia is similar to open source software development, and in both domains, researchers have studied questions about why people participate (WP -- Bryant, Forte, Bruckman 2005; Open Source -- add cite from diss) and the quality of the resulting artifacts (WP -- Stvilia, need to find an Open Source cite). Relatively little work, however, has gone into how these groups coordinate their work. In the case of open source, project roles (owner, developer, bug reporter, technical support) are clearly defined, which likely makes coordination easier (though non-trivial; see Raymond-there's a cite that talks about project ownership transfers if I can find it). In contrast, Wikipedia has few clearly defined roles; those that exist, such as administrator and bureaucrat, are primarily used to grant extra powers such as the ability to block troublesome users from editing and protecting controversial pages from vandalism. Though formal roles are few, Wikipedians recognize a number of informal roles as well, including fighting vandalism, welcoming new users, managing the featured article process (Viegas 2007and writing tools to help the community (Cosley et al. 2007). These informal roles provide open structure is one that provides for legitimate peripheral participation in which new users can begin contributing while observing, and eventually emulating, the behavior of established Wikipedia editors (Bryant, Forte, Bruckman). Studying these informal roles will help us understand how the coordination problem of collective action can be addressed through community processes and monitoring, possibly giving insight into the future health of the community (e.g., if there are not enough answer people in a discussion group, or enough vandal fighters in Wikipedia, the value of the community to others may suffer).

**2.2 Conceptual operationalization of roles**

Social roles can be conceptualized at several different levels of abstraction. The challenge for researchers is to identify roles that matter, roles that help us understand how a social system works. We contend that the best way to identify roles that matter is to begin at the level of meaningful social action, and to work both downward towards identifying the key related behavioral regularities and distinctive positions (signatures of social roles), and upwards to abstract theoretical categories that allow us to tie these particular roles to general research objectives that transcend any particular study or social context.

Starting simply from behavioral regularities or distinctive positions is flawed because it is overly inductive: of the infinite patterns we can find there is no reason to think that the ones that initially stand out will be of any social significance. Starting from abstract categories, like 'altruist', commits the error of over-deduction: simply because we can assign a label to an activity (altruistic) does not necessarily mean that those behaviors are motivated by altruism. The crux is to connect higher levels down to easily observable behavioral regularities and distinctive network positions through higher levels that refer to types of actors that are socially meaningful to participants and relevant to theoretical questions in social research.

The key step is to begin from a role schema for a particular role that fleshes out the distinctive attributes of the role. A role schema is a generalized, abstract mental map that participants in a social setting develop of what a role is, what it means to hold that role, and the wide array of conditional rules about how role holders ought to act, under given range of meaningful conditions. At this level a role is a highly distinctive combination of meaningful, situated actions associated with actors in particular positions. The richness of the map narrows the infinite array of the behavioral and positional attributes

that we could measure, and points to a set of abstract categories through which the particular study can be integrated into a wider body of research.

For instance, in this study we began with a qualitative exploration of the community of Wikipedians as expressed on user pages, community resources, policy discussions, and talk pages associated with articles, user pages, Wikiprojects, and other community related pages. From this, we learned that a large and organizationally important class of Wikipedian is the vandal fighter (counter vandalism editor). The table below indicates expressions of this social role at difference levels of analysis.

**Table 2.2.1 Roles across levels of analysis**

| Components | Indications for example: Vandal Fighter |
|---|---|
| Behavioral Regularities | *Numerous reversions of vandalized pages* |
| Distinctive Nework Position | *Weak connections to high proportions of relative isolates in user talk network* |
| Structure of social action | *Combination of content revisions, sanctions, and efforts to integrate norm violators* |
| Self Identification | *Self-avowel; membership in role related projects and groups* |
| Socially Recognized Types | *Vandal Fighter* |
| Abstract Categories | *Altruist; enforcer of social norms; technical editor* |

## 2.3 Structural Signatures of Social Roles

This paper uses systematic patterns in contribution to identify the signatures of particular roles in Wikipedia. The strategy of using structural signatures social attributes of actors has been applied to a variety of settings. Researchers at Bell Labs identified fraudulent telephone accounts by leveraging patterns in volume, timing, and the identity of in-bound callers (Cortes et al 2000; Cortes and Pregibon 2004). In a similar fashion, consistent driving records are used as proxy for good credit risks where credit data are unavailable (Guseva and Rona-Tas 2001). In more direct measurement situation, Ebay sellers can be identified as reliable or trustworthy through the accumulation of many highly evaluated transactions, a type of emergent, and hard to fake reputation (Friedman and Resnick 2000, Resnick and Zeckhauser 2002). Viegas and Smith (2004) introduce the authorlines visualization and use it to distinguish between different types of online contributors like spammers, debaters, and answer people based on the number and type of messages they contribute to different discussion threads. Our strategy draws on these works in general. In particular we follow earlier studies of roles in the Usenet that used visualizations of cumulative patterns in message contribution as well as attributes of local social networks, like the distribution of neighbors degree to characterize types of discussion spaces (Fisher, Smith, and Welser 2006) and to identify social roles like that of answer person (Welser, Gleave, Fisher and Smith 2007).

## 2.4 Socially meaningful roles in Wikipedia

While the use of Wikipedia is ubiquitous, few users of Wikipedia are aware of the community and resources beyond the articles that provide the social context through which people develop social roles and their careers, and through which newer members learn of the range of existing roles. Different types of pages are organized into different "namespaces" which represent broad classes of activity. Throughout the different namespaces of Wikipedia, there are a range of niche roles that have formed. We focus on substantive experts, technical editors, social networkers and counter vandalism editors. Some users can often be found filling several roles but others choose to focus on a specific activity for the majority of their time. The next section provides a brief discussion of range of socially meaningful roles in Wikipedia and relates them to the organizational challenge of coordinating contribution in the absence of explicit top down management.

**Substantive experts** Substantive experts display extensive knowledge in a topic often citing real world credentials on their user pages. They contribute substantially to pages within a particular subject-area and resolve article related disputes on article talk pages. Though their credentials may or may not come into play, substantive editors are often people who invest time in fact checking and article talk to discuss details of articles.

**Technical editors** There are dozens of areas in Wikipedia where small errors can crop up: spelling, grammar, hyperlink format and out of date facts. The term technical editor refers to all contributors who engage largely in these sorts of incremental improvements and maintenance of Wikipedia's content. Technical editors were identified from participation in technical editing Wikiprojects related to grammar, use of proper copyright information for images, copy editing,

**Counter vandalism** There are editors that spend most of their time trying to find and correct vandalized articles. Users tend to self identify as part of several groups such as the Counter Vandalism Unit; a large percentage of these users agree with the idea and will revert occasional vandalism. The percentage of those users who actually participate in significant anti vandalism efforts is smaller. Though there is a range in numbers of anti vandalism contributions for editors, many editors are devoted strictly to anti vandalism tactics; these were the users included in this directed sample. Because these editors find vandalism either on the Recent changes page, anti vandalism bots or by tracking specific users, counter vandalism editors will have a higher percentage of article edits with little or no relation between article topics.

**Social networkers** Lacking well-defined rules and boundaries, Wikipedia offers users many possibilities for interacting with one another. Those contributors who make excessive use of Wikipedia's networking potential will be referred to as social networkers. Social networkers build strong ties with other users through channels other than article collaboration. They utilize User Talk extensively, make "Wikifriends," and create elaborate profiles that showcase their Wikipedia personalities. Their User Pages are often cluttered with Userboxes, small snippets of self-identifying information including interests, group membership, and personal characteristics. Additionally, social networkers often participate in projects that can be seen as community-building. These include "The Birthday Committee," a variety of projects associated with "Wikipedia Culture," and parts of what was formerly "Esperanza."

## 3. Methods

### 3.1 Analysis strategy

Following (Welser, Gleave, Fisher, and Smith 2007) our analysis proceeds through two stages, an exploratory stage where we use data visualization, descriptive statistics, and content based fact checking to learn the structural signatures of different social roles in Wikipedia. The investigations begin with broad qualitative explorations that identify individuals performing interesting roles, after which maps of the structural positions of populations are used to identify network patterns that differentiate users. The next step is to analyze the context of participation and the content of behaviors of the actors whose interactions formed those social network structures. This iterative process moves between content and structure to refine our understanding of social roles and validate the relationship between structural attributes and behaviors.The remainder of the methods section includes a description of our population samples, our data, wikipedia namespaces as they pertain to social roles,

### 3.2 Samples

**Directed** The directed sample includes hand picked contributors who, based on the types and content of their edits, as well as their userpages seem to perform the roles of substantive expert, technical editor, counter vandalism, or social networker. We use this sample to document the patterns and exceptions we find within each type and to develop insight into possible metrics for distinguishing these roles from other types of editors. For technical editors and counter vandalism we were able to use lists of participants in related Wiki projects to identify possible role players.

**Cohort** The cohort sample collects all editors who created accounts during the month of January 2005 and tracks their edits to namespace across their subsequent career, to the limits of our data (about 20 months). This sample allows us to measure the proportion of new entries that fall into each of the observed roles. Unfortunately this sampling frame will not include most of the dedicated wikipedians, many of whom began editing prior to our cohort sampling frame, hence the lagged sample.

***Dedicated Editor*** This sample highlights dedicated contributors by generating a list of all contributors who began editing during Jan 2004 or earlier, and also made at least one edit during January of 2005. The same monthly edit space and structural data as for cohort sample is collected for this sample.

### 3.3 Data

Our data are drawn from records of edits, organized by editor, distributed in the Fall 2006 data dump, available for free download from MediaWiki (http://meta.wikimedia.org/wiki/Data_dumps). Data extends from the beginning of the English Wikipedia to October of 2006. We parsed the data into sub-files that allowed us to identify users who were active from particular time periods and collect monthly counts of the edits to different types of pages within Wikipedia. The Wikipedia file system is subdivided into "namespaces" which represent general categories of pages based on their function (see http://en.wikipedia.org/wiki/Wikipedia:Namespace).

**Table 3.2.1: Coded Namespace Types**

| Category | Namespaces | Notes |
|---|---|---|
| Content | Article {0}<br>Image {6} | Creation and editing of content are direct acts that contribute to the writing of the encyclopedia. |
| Content Talk | Article_talk {1}<br>Image_talk {7} | Content talk pages represent discussion of changes, negotiation of proper direction for particular pages, sanctioning of norm violators and other management tasks directly related to the content of the encyclopedia. |
| User | User {2} | Edits to user pages represent two general activities: edits to ones own user page involve the presentation of self; edits to User pages of others often indicate managerial tasks like banning accounts of repeat vandals. |
| User Talk | User_talk {3} | Edits to usertalk pages are good indications of direct person to person communication in Wikipedia. These directed ties are made for a variety of reasons, including communication among friends, negotiation of proper edits, issuing thanks or public recognition (barnstars), sanctioning of norm violators, etc. Edits here are roughly analagous to making comments on personal blogs or replying in a Usenet discussion. |
| Wikipedia | Wikipedia {4}<br>Wikipedia_talk {5} | Edits to Wikipedia and Wikipeida Talk pages indicate involvement with community aspects of Wikipedia. Example pages include: community portal, help desk, village pump. |
| Other / Infrastructure | MediaWiki {8,9}<br>Template {10,11}<br>Help {12,13}<br>Category {14,15}<br>Portal {100,101} | The most commonly edited of the remaining namespaces are primarily concerned with the construction of infrastructure that facilitates the work of others. These include things like the creation and editing of templates, categories, media wiki, portals, and help pages. The associated talk pages are included, and are indicated with the second namespace number in brackets { }, at left. |

Edits to different namespaces often reflect different types of tasks. As we discuss in detail below, the proportion of edits dedicated to a particular namespace as well as the distribution of those edits across time and across pages can reveal signatures of different social roles. In the following table we have divided the edits into six areas that represent different types of contributions to Wikipedia. We use histograms to compare the distribution of edits between namespaces and local network visualizations to graph structure of relationships in the user talk namespace.

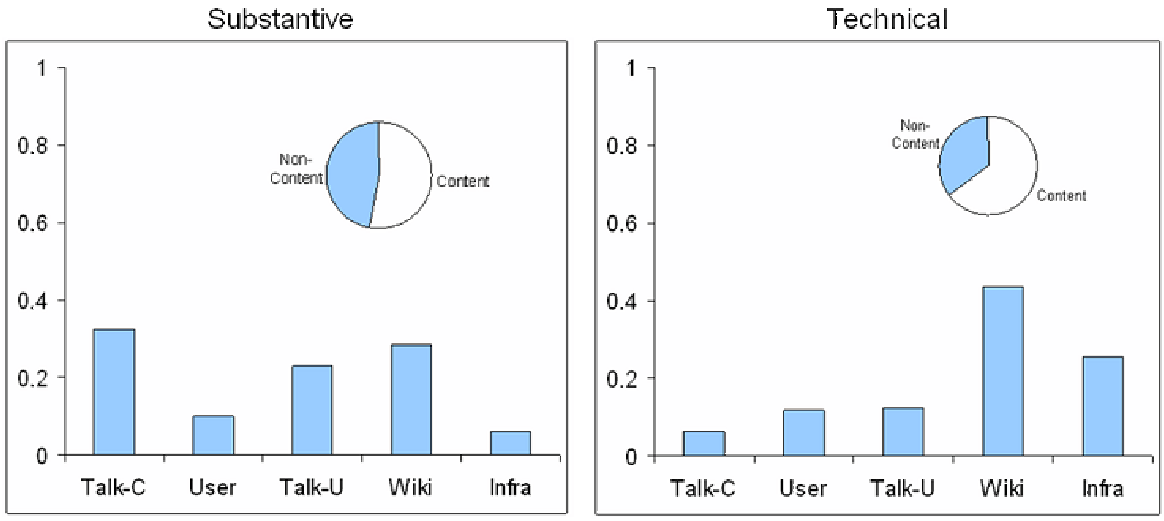
**Table 3.2.2: Sample comparisons for edit distributions**

| | | | | Directed Sample Types | | | |
|---|---|---|---|---|---|---|---|
| | **Cohort** | **Dedicated** | **Directed** | *Substantive* | *Technical* | *C-Vandalism* | *S-Networkers* |
| Months Active | 4.53 | 31.64 | 16.08 | 19.92 | 19.80 | 12.20 | 9.71 |
| Total Edits | 251.11 | 5059.80 | 7462.82 | 7465.25 | 12142.00 | 6309.90 | 2421.14 |
| Edit Rate | 55.43 | 159.91 | 464.19 | 374.82 | 613.23 | 517.20 | 249.24 |
| Content | 68% | 73% | 59% | 53% | 65% | 68% | 21% |
| Content Talk | 9% | 8% | 7% | 15% | 2% | 4% | 6% |
| User | 5% | 3% | 6% | 5% | 4% | 5% | 22% |
| User Talk | 6% | 5% | 8% | 11% | 4% | 8% | 26% |
| Wikipedia | 9% | 9% | 14% | 13% | 15% | 11% | 23% |
| Infrastructure | 3% | 2% | 6% | 3% | 9% | 4% | 1% |

## 4. Analysis

**4.1 Distribution of edits** How are different roles revealed by how people divided their edits across namespace categories? We begin to address this question by comparing the average distributions of edits across the six namespace areas discussed above, focusing on the four role types in our directed sample. The pie chart represents the proportion of all edits devoted to content. The histogram shows how the non-content edits are divided among content talk, user, user talk, wikipedia, and other. The following figures depict averages taken from the directed sample.

Figure 4.1.1

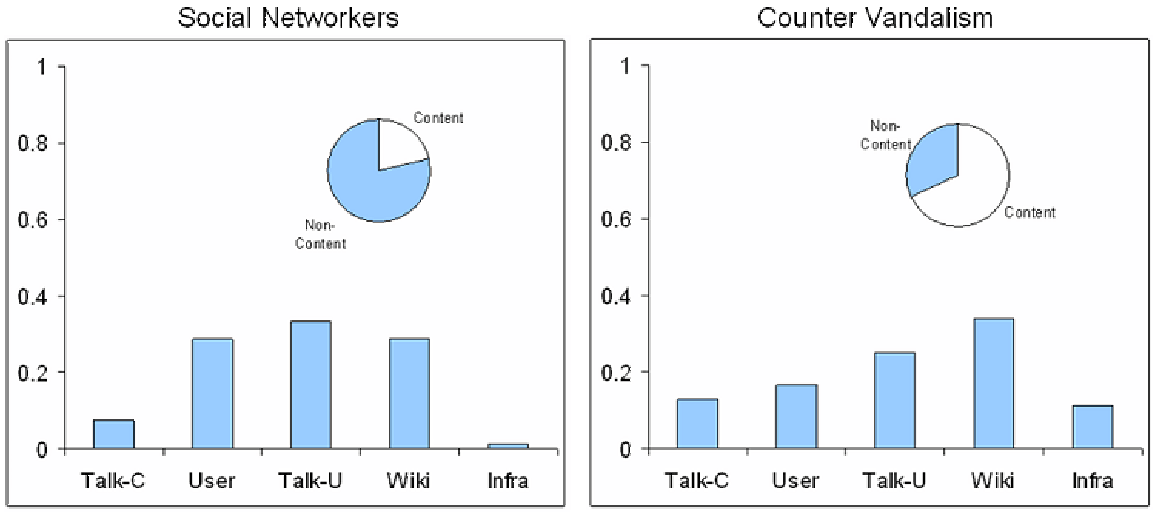



The average **substantive expert** makes only about 50% of edits to content space. This rate is about 10% lower than what we observed for anti-vandals or technical editors. However, substantive experts have a much higher rate of posting to content talk. This suggests generally, that when substantive experts contribute to content pages, their contributions are likely to be more costly (take more time and thought, and are more likely to require explanation, justification and discussion on the content talk pages). Their distinctive investment patterns are defined by the combination of a lower content edit rate with elevated content talk, followed by edits to Wikipedia namespace pages. The remaining areas being low.

**Technical editors** make numerous small changes to content pages, frequently specializing in a particular type of problem (spelling, grammar, faulty links, improper copyright information, etc.) thus there are several subtypes of specific edit patterns within this class of editors. However, within this

class there is a shared tendency to invest primarily in content edits, with contributions to Wikipedia pages a strong second in edit rates. These editors are primarily making the specific content changes associated with the issue(s) they like to address and holding discussions about community and infrastructure related to reinforcing those tasks (hence the Wikipedia name space edits).

Figure 4.1.2




**Counter vandalism editors** achieve edit profiles very similar to other types of technical editors. They have moderately high rates of content edits, followed by investment in Wikipedia and user pages. Where technical editors concentrate their non content edits primarily in Wikipedia namespace. We were initially surprised to see that counter vandalism editors had such high rates of edits to the User and User Talk namespaces. However, inspection of authorlines for counter vandalism editors showed that large portions of their User page edits were made to *other* people's User pages. This somewhat unusual behavior is driven by counter vandalism editors that are also admins, and in the course of blocking a vandal they make an edit to that user's page. Counter vandalism work also often involves posting warnings on user talk pages, which explains the relative frequency of edits to user talk pages compared to other technical editors. With the exception of these highly specific edits to the user pages of banned vandals, the edit distribution of counter vandalism editors seems indistinguishable from other types of technical editors.

Finally, the **social networking editors** provide a sharp contrast to the other three types—they invest very little in content edits, and invest primarily in their own user page. Their second priority involves user talk, and next are investments in Wikipedia namespace pages, which are often associated with community building and support for social interaction among editors. Finally, their investment in infrastructure, as represented by the other category is negligible.

In depth investigations into role types

## 4.2 Network structure

Patterns in relationships can help distinguish between roles in online spaces. The following discussion describes differences in the structure of relationships in the user talk network, comparing examples from our directed samples of substantive experts and technical editors. These graphs display patterns of connections between ego, all one degree neighbors of ego, and the ties between those alters (omitting all ties to alters degree 2 or greater from ego). The arrow travels from author to person whose user talk page received the edit, red arrows indicate that ego was the author, blue that an alter was the author.

Substantive Experts display extensive knowledge of particular topics, contributing a great deal of content in that subject-area and resolving content-related disputes on the "talk pages" associated with

each Wikipedia content page. Technical Editors enforce the formatting standards of Wikipedia while policing grammar, style, spelling, hyperlink accuracy, and other minor issues. These tasks are related, but many Wikipedia contributors specialize, developing their expertise in ways which cause them to distribute their contributions differently. These two social roles are of particular interest, both because they are crucial to the core task of the encyclopedia and because they possess clearly contrasting structural signatures.

Technical Editors, like copy editors, have a repetitive task that is largely independent of the content of any given page. Thus, such Wikipedia contributors should have brief interactions with many other participants, serially moving in and out of relationships with Substantive Experts who have a more focused and durable connection with a smaller set of pages. Substantive Experts, by contrast, are more likely to engage in ongoing conversations with other Substantive Experts as they work to improve pages in their shared areas of expertise. Therefore, the social network structures that surround these social roles may illustrate the distinctions between them. The network structures of Technical Editors should feature outward connections to alters who are otherwise not connected to each other (resembling the Answer Person in Usenet). The performance of the role of Substantive Expert should generate a dense interconnected network (resembling the Discussion Person).

**Figure 4.2.1: Local Networks for Technical Editors and Substantive Experts**

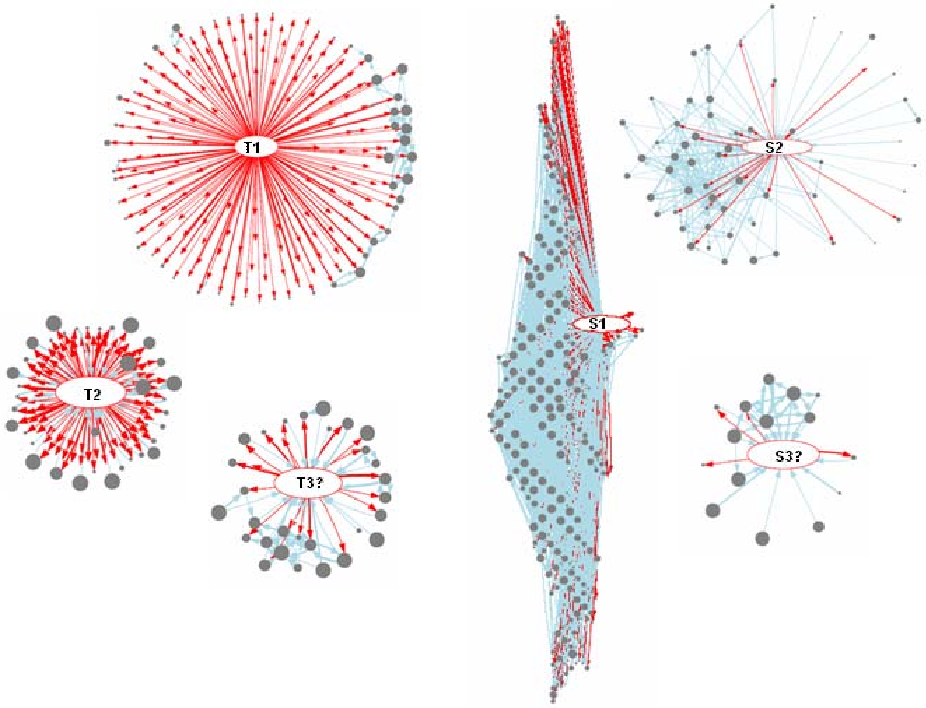



Figure 2 depicts visualizations for the local social networks surrounding three Technical Editors (T1-3) and three Substantive Experts (S1-3). A directed outward tie indicates that an Editor posted a message to another Editor's "user talk page", a page on which conversations associated with each Wikipedia participant are stored. A reciprocating tie from another Editor is indicated with a lighter gray line. The size of each node indicates the measure of the out-degree from that Editor.

Technical Editors tend to have more outward ties to other Editors who are otherwise unconnected to one another. The illustration of Technical Editor 1 (T1) in Figure 2 best fits this ideal model — displaying dozens of outward ties and very few connections among those alters. The other two Technical Editors displayed here have some deviations from the expected structural pattern. T2 is still fairly close to the ideal type, but does have a higher in-degree and a large number of connections to nodes with high out-degree, while the ideal type would have lower in-degree and fewer connections to such socially active editors. T3 differs dramatically from both T1 and T2, with a relatively highly

interconnected local network. T3 is almost more similar to the Substantive Experts, who are more likely to have dense interconnections in their local networks. The substantive experts have more variance in the number of alters and the density of their local networks than the technical editors.

### 4.3 Distribution of roles within cohort and dedicated samples

**** this analysis step still needs to happen *****

-- assign members of cohort and directed sample to role types according to indicators discussed above and report relative frequency.

## 5. Discussion

Our discussion considers two general themes of relevance to scholars of Wikipedia in particular and scholars of organizations and of online community more generally. First, In our discussion we explore the possibility that the key to solving the coordination problem of Wikipedia authorship depends on the solution to a prior problem, the challenge for each author to find their role in Wikipedia. We consider how entry rates into, and exit rates out of different roles may place limits on the growth/quality of Wikipedia, and consider some institutional affordance for aiding the function and growth of distributed collaborative systems like Wikipedia. Second, we consider how different affordances of online social systems are likely to affect role adoption, and therefore how those patterns of role adoption will affect the growth of the community.

Wikipedia provides a rich array of ways to participate and a complex set of interrelated tasks, all of which are reveled gradually to contributors as they develop their expertise. Compared to early forms of web community (like Netscan and Web forums) Wikipedia is more complex, nuanced, and user-modifiable. Because it seems to provide more room and more ways for people to develop we might expect Wikipedia to be better able to retain participants once they get past initial levels of participation, but that it might be harder for people to initially enter the activity if those options for participation are not apparent to most users. This suggests two important dimensions of design that may affect length and nature of participant careers: the complexity of the behavioral terrain and the accessibility of that range to participants.

Maybe something about the roles not being exhaustive and not necessarily mutually exclusive?... The roles admins may play in giving more structure?

Dependency of the whole system on the particular roles. sick in absence of certain role, interdependence. Organic solidarity. . . Functionalism explanation of crime-- need for other roles. similar in a way to "in praise of dead wood thesis"

**References**

Brandes, U., Erlebach, T. ed. 2005. *Network Analysis: Methodological Foundations*. Springer.

Brandes, U., Raab, J., Wagner, D. 2001. "Exploratory Network Visualization: Simultaneous Display of Actor Status and Connections." *Journal of Social Structure*. 2 (4).

Burt, R. S. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, Mass.: Harvard University Press.

Burt, R. S. 2004. "Structural Holes and Good Ideas." *American Journal of Sociology* 110:349-399.

Callero, P. L. 1994. "From Role-Playing to Role-Using - Understanding Role as Resource." *Social Psychology Quarterly* 57:228-243.

Chesney, Thomas. 2006. An empirical examination of Wikipedia's credibility. First Monday 11-6.

Coleman, J. S. 1990. *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.

Cortes, C., Pregibon, D. 2001. "Signature-Based Methods for Data Streams." *Data Mining and Knowledge Discovery* 5:167-182.

Cosley, Dan, Dan Frankowski, Loren G. Terveen and John Riedl. 2007. SuggestBot: using intelligent task routing to help people find work in Wikipedia. Proceedings of the 12th International conference on Intelligent User Interfaces (IUI 2007), pages 32—41.

Donath, J. S. 1996. "Identity and Deception in the Virtual Community." In *Communities in Cyberspace*, pp. 29-59. Edited by P. Kollock & M. Smith, London: Routledge.

Donath, J. S. 1999. "Visualizing Conversation." *Journal of Computer-Mediated Communication,* 4(4).

Erickson, T., Herring, S. 2005. "Persistent Conversation: A Dialog Between Research and Design." *Proceedings of the 38th Hawaii International Conference on System Sciences (HICSS 2005).* IEEE Press.

Faust, K., Skvoretz, J.. 2002. "Comparing Networks Across Space and Time, Size and Species." *Sociological Methodology.* 32: 267-299.

Fehr, E., Fischbacher, U. 2003. "The Nature of Human Altruism." *Nature*. 425: 785–791.

Freeman, L. C. 2000. Visualizing Social Networks. *Journal of Social Structure*, 1(1).

Friedman, Eric and Paul Resnick. 2000. The Social Cost of Cheap Pseudonyms. *Journal of Economics and Management Strategy* 10(2): 173-199.

Fisher, D., Smith, M., Welser, H. T. 2006. "You Are Who You Talk To: Detecting Roles in Usenet Groups." In *Proceedings of the 39th Hawaii International Conference on Systems Sciences (HICSS)*, January 4-7. Computer Society Press.

Giles

Golder, S. A. 2003. *A Typology of Social Roles in Usenet.* Unpublished Senior Honors, Harvard University, Cambridge, MA.

Guseva, A., and Rona-Tas, A. 2001. "Uncertainty, Risk, and Trust: Russian and American Credit Card Markets Compared." *American Sociological Review* 66:623-646.

Haythornthwaite, C., Hagar, C. (2005). "The Social World of the Web." *Annual Review of Information Science and Technology 39*: 311-346.

Herring, S. C. 2004. "Slouching Toward the Ordinary: Current Trends in Computer- Mediated Communication." *New Media & Society* 6(1): 26-36

Lampe, C., Resnick, P. 2004. "Slash(dot) and Burn: Distributed Moderation in a Large Online Conversation Space." CHI Vienna, Austria.

Merton, R. K. 1968. *Social Theory and Social Structure*. New York,: Free Press.

Nadel, S. F. 1964. *Theory of Social Structure.* New York: Macmillan

Neustaedter, C., Brush, A. J., Smith, M., Fisher, D. 2005. *The Social Network and Relationship Finder: Social Sorting for Email Triage.* CEAS.

Olson, M. 1971. *The Logic of Collective Action; Public Goods and the Theory of Groups*. New York: Schocken Books.

O'Madadhain, J., Fisher, D., White, S., Boey, Y. 2003. "The JUNG (Java Universal Network/Graph) Framework." *Technical Report UCI-ICS 03-17.* School of Information and Computer Science, University of California, Irvine.

Ostrom, E. 1991. *Governing the Commons: The Evolution of Institutions for Collective Action*. New York: Cambridge University Press.

Parsons, T. 1951. *The Social System.* Glencoe, Ill.: Free Press.

Pfaffenberger, B. 2003. "A Standing Wave in the Web of Our Communications: Usenet and the Socio-Technical Construction of Cyberspace Values." In *From Usenet to Cowebs.* Edited by C. Lueg & D. Fisher, pp. 20-43. London: Springer Verlag.

Read, Brock.

Resnick, P. and Zeckhauser, R. Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System. The Economics of the Internet and E-Commerce. Michael R. Baye,

editor. Volume 11 of Advances in Applied Microeconomics. Amsterdam, Elsevier Science. 2002. pp. 127-157.

Sack, W. 2002. "What Does a Very Large-Scale Conversation Look Like? Artificial Dialectics and the Graphical Summarization of Large Volumes of E-mail." *Leonardo* 35(4): 417-426

Skvoretz, J., Faust, K. 2002. "Relations, Species, and Network Structure." *Journal of Social Structure* 3.

Smith, M. A. 2003. Measures and Maps of Usenet. In *From Usenet to Cowebs.* Edited by C. Lueg & D. Fisher, pp. 47-78. London: Springer Verlag.

Turner, T.C., Smith M., Fisher, D., Welser H. T. 2005. "Picturing Usenet: Mapping Computer-Mediated Collective Action." *Journal of Computer Mediated Communication*. 10(4).

Viegas, F. B., Smith, M. A. 2004. *Newsgroup Crowds and Author Lines: Visualizing the Activity of Individuals in Conversational Cyberspaces*, Big Island, Hawaii: IEEE.

Wasserman, S., Faust, K. 1994. *Social Network Analysis*. Cambridge: Cambridge University Press.

Weber, S. 2004. *The Success of Open Source*. Cambridge MA. Harvard University Press.

Wellman, B. 2001. Computer Networks as Social Networks. *Science, 293*(5537), 2031-2034.

Welser, Howard T., Eric Gleave, Danyel Fisher, and Marc Smith. 2007. "Visualizing the Signatures of Social Roles in Online Discussion Groups." *The Journal of Social Structure.* 8(2).

Wenger, E. 1999. *Communities of Practice: Learning, Meaning, and Identity*. Cambridge, England: Cambridge University Press.

White, H. C., Boorman, S. A., Ronald L. Breiger. R. L. 1976. Social-Structure from Multiple Networks .1. Blockmodels of Roles and Positions. *American Journal of Sociology, 81*(4), 730-780. 0

Winship, C. 1988. "Thoughts About Roles and Relations - an Old Document Revisited." *Social Networks* 10:209-231.