Reflections and Reactions to Social Accounting Meta-Data

Eric Gleave[1] and Marc Smith[2]

[1] University of Washington, USA
[2] Microsoft Research, USA

1. Introduction

Online systems are becoming increasingly social environments in which people share advice and experiences in threaded discussions, photos, videos and other files in systems like Flickr and YouTube, and display details of their social lives through a host of social networking sites. Yet even as these settings provide rich content, that content does not automatically provide us with social cues that can reveal what an interaction might mean, who we are interacting with, or the nature of their underlying character. As more of social life is embedded in these systems, we come to want and need systems for expressing identity and building reputations that can help us resolve some of this uncertainty. Because interaction in these settings leaves traces, however, we can look at histories and patterns of actions from hundreds of interactions. These types of accumulated reputations can reveal a great deal.

To varying extents online social spaces can be designed to reflect behavioral histories to users. When such systems are made available, how are they used by participants? The Netscan system generates just such patterns of author and newsgroup activity for Usenet and publishes them to the web. In this paper, we explore the ways

the data generated by Netscan is used by the participants themselves as evidenced by the messages posted to Usenet in which the Netscan system is mentioned by the string "netscan.research", a fragment of the system's URL.

Anonymity and reputation ambiguity can be alienating to new users unfamiliar with a particular community [4, 24]. This can lead to the domination by socially disruptive individuals who are otherwise fairly rare [19]. The availability of social accounting meta-data can reduce or eliminate this barrier to the formation of durable communities. Such information permits the establishment and maintenance of norms and the identification of the reputations of unfamiliar participants [4]; facilitates boundary maintenance for communities [9]; and reflects the character of groups to potential new members [7, 18, 24]. Provision of such surveillance is a double edged sword, however. While it provides these advantages to the construction and maintenance of communities, it also comes with the tinge of Big Brother, NSA wire-tapping, and a sense of being watched and monitored without explicit consent [1].

How populations react to the public availability of social accounting meta-data acquired through surveillance and analysis of already publicly available data remains an open question. How do participants react to these reputation systems? More importantly for the establishment of communities, how does such a system influence participants' behavior? On balance, do such tools increase the presence and vitality of community or do they frighten users away? These are increasingly relevant questions as digital traces become ubiquitous and data warehousing, transfer, and presentation become cheaper, easier, and more common.
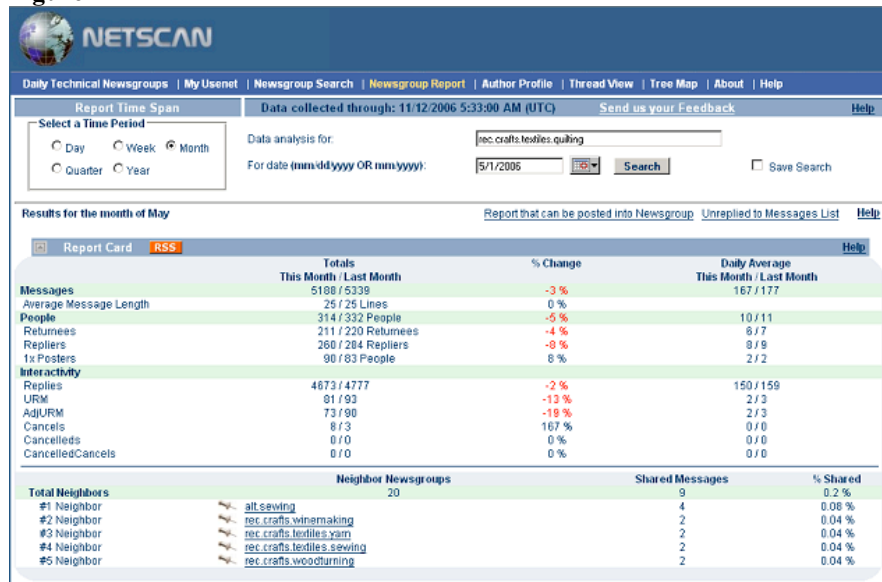
These questions are not particular to research in online communities. The communities developed in newsgroups can be viewed as instances of other standard social processes of interest. Social accounting data usage lends understanding to processes of group and community formation [14]. This process itself involves the development and enforcement of group norms [6, 15, 17]. Interactions invoking social accounting data tell us about how individuals identify characteristics of others in anonymous contexts [8, 20, 16, 25] as well as how trust is established [7, 13, 11, 22]. While the context of Usenet and online community may not be perfectly analogous to real

world, face-to-face society, the fundamental aspects of these processes are present and are rendered easier to study.

This paper investigates these questions in the realm of Usenet via the Netscan Project. Netscan is a data collection, aggregation, analysis and reporting project that collects messages from Usenet and related repositories of threaded conversations. Since late 1999, Netscan has collected all data associated with the headers of messages, including subject, time of post, reply relationships, the newsgroup(s) it was posted to, etc., and recently includes the content of messages as well. The system then reports meta-data and summaries through a web interface that allows users to search for information on authors, threads, and newsgroups at time ranges from one day to one year.

The Usenet is a large system for threaded discussion, involving over a hundred thousand topically differentiated newsgroups (for a general description of the Usenet, see [21]). Usenet messages are distributed through news-servers provided by Internet Service Providers, dedicated software clients (newsbrowsers), and web sites like Google Groups (http://groups.google.com).

Figure 1

NETSCAN

Daily Technical Newsgroups | My Usenet | Newsgroup Search | Newsgroup Report | Author Profile | Thread View | Tree Map | About | Help

| Report Time Span | Data collected through: 11/12/2006 5:33:00 AM (UTC) | Send us your Feedback | Help |

Select a Time Period
- Day - Week - Month
- Quarter - Year

Data analysis for: rec.crafts.textiles.quilting
For date (mm/dd/yyyy OR mm/yyyy): 5/1/2006   Search   Save Search

Results for the month of May   Report that can be posted into Newsgroup   Unreplied to Messages List   Help

Report Card RSS   Help

	Totals This Month / Last Month	% Change	Daily Average This Month / Last Month
Messages	5188 / 5339	-3 %	167 / 177
Average Message Length	25 / 25 Lines	0 %	
People	314 / 332 People	-5 %	10 / 11
Returnees	211 / 220 Returnees	-4 %	6 / 7
Repliers	260 / 284 Repliers	-8 %	8 / 9
1x Posters	90 / 83 People	8 %	2 / 2
Interactivity			
Replies	4673 / 4777	-2 %	150 / 159
URM	81 / 93	-13 %	2 / 3
AdjURM	73 / 90	-19 %	2 / 3
Cancels	8 / 3	167 %	0 / 0
Cancelleds	0 / 0	0 %	0 / 0
CancelledCancels	0 / 0	0 %	0 / 0

	Neighbor Newsgroups	Shared Messages	% Shared
Total Neighbors	20	9	0.2 %
#1 Neighbor	alt.sewing	4	0.08 %
#2 Neighbor	rec.crafts.winemaking	2	0.04 %
#3 Neighbor	rec.crafts.textiles.yarn	2	0.04 %
#4 Neighbor	rec.crafts.textiles.sewing	2	0.04 %
#5 Neighbor	rec.crafts.woodturning	2	0.04 %

Because Netscan has existed for several years and some Usenet authors are aware of it (as evidenced by the over 250,000 unique visitors per year) we should wonder how the interface may affect interaction and self-perceptions within Usenet. Newsgroup Report cards are one especially relevant source of information for Usenet authors.

The top section of the report, depicted in figure 1, includes aggregated statistics about the newsgroup such as message length; counts of posters, returnees, and first time posters; and measures of interactions such as number of reply messages. Below this are lists of the top forty threads (by message count) and top forty authors (by days active in the selected time period). This report card serves as a detailed portrait of a newsgroup in terms individuals, population, and conversations.

2. Data and Methods

By searching for the string "netscan.research" though Google Groups Usenet search index system, we collected the contents of every message that explicitly references the Netscan project. Removing those that quoted previous messages with the string, this resulted in 4,939 unique messages posted between April 29, 1999, and November 5, 2004. The full threads containing these messages were also extracted. This allows a coder to see the full context of a message rather than just the message itself, as the same words in one newsgroup may carry a completely different meaning in another. The messages come from 1,241 unique newsgroups.

These message bodies were then placed into a custom-built analysis tool implemented in the Microsoft Access database application. This tool presents the analyst with a message body, its thread of origin, and subject line. A coder can apply any number of pre-authored tags to the message to label it. In this study labels were applied to indicate the ways Netscan data was presented and the nature of any reactions to the data presented in a message. A coder can also create new tags if the existing tags do not capture the observed behavior.

Messages were presented to the content analyst in random order and initially no tags were defined. Data presented in this analysis

comes from the coding of an initial sample of 1,454 messages, about 30% of the total data set. Along with the types of data used and affective reactions, messages were also coded for what language they were in whenever the predominant language used was not English. Of the tagged sample, 34.5% were written in a language other than English, including 17 identified languages and 11 messages in unidentified languages. As accuracy of interpretation cannot be determined for messages written in languages the coder was not fluent in, these messages are disregarded here, leaving 952 messages in this analysis.

The following is an example of the types of tags associated with a typical message. This message was coded with the tags: self-competitive, self-bragging, scary, and newsgroup reputation.

> Woo-hoo I'm No. 4, but they have a scary amount of info on me! Perhaps more worryingly, our nearest "neighbour" is alt.sex.fetish.boyfeet...

3. Analysis and Results

The tagging schema established dynamically in the process of the content analysis is comprised of four general themes: Group, Self, Other, and Netscan. These are first explained in turn and then described quantitatively as found in the data. The proportions of uses can be found in Table 1. Note than any one message can contain any number of usages or combination of usages. As such, the proportions presented in Table 1 do not sum to 1 nor do the counts sum to the total N.

Group messages include any reference to a specific newsgroup. The usages include concern for the health of a newsgroup, comparison of one newsgroup to another, inclusion of Netscan data in a newsgroup's FAQ, concern for the reputation or cross posting tendencies of a newsgroup, links to the Netscan data on a particular newsgroup, or direct posting of a newsgroup's Netscan Report card. Such usage is typical in messages that are part of an initial introduction of a newsgroup to the existence of Netscan as a way of encouraging other readers to look at the data provided through the Netscan website.

Self messages specifically reference the author of the post. These usages include linking to the author's profile in Netscan, references to the cross posting statistics, the claiming of bragging rights based on some Netscan measurement of author activity (i.e. "I am most active"), expression of competitiveness, concern for reputation, and referencing Netscan data in defense of the author in discussions that call the author's authority or reputation into question. Initial messages about Netscan in a newsgroup often also include the author's reactions to information about their own activity as presented by the system, but instances of self references more commonly appear as part of the culture of a newsgroup. These are predominantly competitive in nature.

Table 1. Tag Frequencies in Messages and Categories

	Proportions Within:	
Tags	Messages	Category
Group (N=490)	0.51	
Link		0.81
Health		0.17
Comparison		0.06
Cross posting		0.04
Reputation		0.02
Self (N=206)	0.22	
competitive		0.49
bragging		0.40
reputation		0.21
Others (N=294)	0.31	
Reputation		0.54
Competition		0.41
Troll Check		0.05
Netscan (N=458)	0.48	
Features/bugs		0.47
Affect		0.54
Affect (N=246)	0.26	
Cool/Interesting		0.58
Big Brother		0.28
Scary		0.19
Bad		0.10
"MS is Evil"		0.08
Note: N=952		

Other messages reference information about specific posters other than the message author. These usages are similar to those under *self* but focus on another poster. Most instances concern the reputation of others or foster competitiveness between newsgroup participants, shifting the focus of competition from individuals to the newsgroup as a whole.

Netscan messages talk about the interface, provide a link to the website, or express an affective response to the tools. The majority of instances are comments on the features of the tools or data. Most common among these are complaints that the rankings of authors are based on number of days active rather than another readily available metric such as threads initiated, total posts, total replies, or a measure of the volume of contribution in bytes or lines. The other large category within *Netscan* tags are emotional reactions. A clear majority of these references see the tools and data as "cool" or "interesting". A distant second to these positive reactions are negative ones finding the tools to be 'bad', 'scary', or reminiscent of 'Big Brother'. Also included in this schema are explicit negative affect towards the Microsoft Corporation.

Groups and Community

Over half of all messages include a reference to the newsgroup, making this category the most commonly identified. Within this, 80% contain a link to the Netscan data on the newsgroup. This basic gesture appears in most initial messages a newsgroup receives concerning Netscan and allows users to investigate the website themselves. This prevalence is also due to the tendency for self-identified leaders of newsgroups to incorporate Netscan data into periodic messages that describe the purpose and history of the newsgroup, whether these are formalized announcements and lists of frequently asked questions or informal accounts given by users who notice new Netscan data has been released. The newsgroup that has referenced Netscan the most total number of instances has done so by linking to the website in a FAQ and newsgroup description consistently every month.

Following linking to a newsgroup report card, the next most common usage in this category is expressing a concern for the health

of the newsgroup. Seventeen percent of messages in this category are explicitly concerned about participant turnover, a decline in number of participants, a decline in number of posts, or a perceived reduction in quality of the newsgroup. The following example is typical of this type of message.

> Posts here are down 52% since this time last year. And there is 54% less people posting here now then a year ago as well. I wonder why people dont post here as much these days? [...] Go here for all the stats. Pretty cool site you can use it for all the newsgroups out there: [link to Netscan report for that newsgroup]

Authors of these posts are typically long-term participants who play a role of community elders. Newsgroups which display this behavior are those that have a sense of community and a desire for continuity in membership and resultant norms. Emotional support groups and avocation groups express this concern, while technical support groups were never observed doing so. This is to be expected as technical support groups have a small number of regular participants and little to no discussion among this core.

Less common uses in the category of *group* are worries about cross posting behavior, concern for the possible reputation of the newsgroup, such as the first example above, and comparisons, often competitive, between one newsgroup and another, such as the following excerpt.

> [That group is] not all that dead - I did a netscan.research.microsoft.com report on alt.asian-movies and rec.arts.movies.local.indian, and the Indian group had about 78% of our number of posts, 40% of our number of posters and about the same number of replies that we had.

Each of these indicates that the poster is identifying the newsgroup as a community of likewise interested others. The concern for reputation typically arises from the cross posting behavior of participants. Since Netscan reports the names of newsgroups that are "neighbors" based on frequency of messages that are cross posted (or shared) between them, newsgroups often find themselves associated with newsgroups which the participants view as distasteful. Although these relationships are often the result of cross posted spam messages, their appearance in a newsgroup's report card often leads

to an explicit statement of cross posting norms for that newsgroup
from a regular poster.

Reputation and the Presentation of Self in Internet Life

Twenty-two percent of messages make explicit reference to the
author of the post. Nearly half of these are competitive in nature.
These messages are typically triumphant declarations of the author's
position in the Netscan rankings, based on days active in the period.
This feature causes a lot of returns to the website as positions typi-
cally fluctuate from month to month, allowing this badge of status to
rotate amongst the top several posters in a newsgroup. Furthermore,
40% of self-referencing messages are explicitly bragging about
rankings, such this simple post from a college sports newsgroup:
"I'm number 1!"

   Following these competitive uses, authors write about themselves
out of a concern for their reputations. Many complain that the data
is inaccurate. Since Netscan tracks authors based on their e-mail
address, a single person may be in the database any number of times
based on how many different addresses s/he used. Many seasoned
users change parts of their e-mail addresses in order to avoid collec-
tion by spam bots or receive unintentional direct e-mails. This leads
many users to complain that they have actually been active in a
group longer than Netscan says they have. These authors are mak-
ing a claim to greater seniority than Netscan gives them, and they
are often upset at the discrepancy.

> Hmm, I think 1998 was my big year for this stuff. Of course they can't
> go back that far.

   Authors also use the Netscan data to defend their reputation.
When arguments break out in Usenet, there is usually no way to de-
termine facts about participants. The data available through Net-
scan, however, permit the comparison of long-term reputations.
This process leads authors to invoke Netscan to establish their tenure
in a newsgroup or their participation in a range of newsgroups on a
similar topic for the establishment of expertise. In the reverse of
this, authors defend their reputations by claiming Netscan presents
false information. This is usually because an author is listed as par-

ticipating in newsgroups s/he did not know existed or which have a negative reputation effect.

As might be expected, holding up a mirror that reflects authors' activity changes their behavior [8, 20]. With their posting behaviors and participation patterns laid bare and easily accessible, authors express an intention to change their future behaviors in order to manage their reputations [12]. The heightened sense of awareness of the consequences of cross posting, for instance, causes authors to express a desire to be more careful in this practice in the future. Authors also say they will use only one string to indicate their e-mail address in order to keep all of their behaviors under one name. Though not clearly evident in the data, it is quite likely that authors also intentionally change their address when participating in groups outside their normal range. Minor explicit instances of this appear in couples who share an address through their home computer but, after exposure to Netscan, make sure their addresses are different from one another so as to not collapse the behavior of two or more people into one reputation.

Others: Friends, Foes, Flame Warriors, and Trolls

Netscan is used heavily for the discussion and establishment of the reputation of other posters, with 31% of messages mentioning a specific poster. These instances fall into two generally separate classes: the fostering of competitive behavior in other newsgroup participants and discussion of reputations.

The most common reference to other participants is the invocation of reputation, with 54% of messages that reference others concerning reputations. This behavior is to be expected in a discussion space where the reputation and history practices of face-to-face communication are not accessible [11, 9]. Users are interested in knowing more about the history and viewpoints of the other participants whose content they encounter. Netscan provides the ability to quickly establish the other interests of a fellow participant based on what newsgroups s/he posts to and their patterns of content contribution which can characterize the type of poster s/he is. This message, for example, is part of a larger conversation as to the motives behind one author's posting behaviors.

> ... just for your information, [Poster X] posts to over 17 other newsgroups on Usenet, and [Poster X]'s name is #20 on the list of most active posters to [this group], with [Poster Y] being #1 and I being #2.

The tools are commonly used to make informed guesses as to whether someone is a flame warrior, troll, or spammer (for discussion of these ideal typical roles see [26]). The pattern of posting for each of these Usenet roles is discernable from the author profiles in Netscan. Newsgroups protect themselves against invasion and domination by authors who are disruptive to the normal operation of a particular community. In several instances, users identify a newcomer as a "flame warrior" or potential "troll" (a user who incites flame wars by contributing intentionally provocative topics) early on and encourage other users not to engage with the newcomer, with 4.8% of messages about others explicitly using Netscan to check for trolls. This ability serves as an early warning for a newsgroup and allows the quarantine of these potentially disruptive participants.

References to others in a competitive spirit comprise 41% of all messages that refer to specific authors. As new Netscan data on newsgroups is updated, participants use the opportunity to discover the current rank of the key participants in their newsgroups. The following example, for instance, chastises other members of the newsgroup for not contributing as much as this author while specifically encouraging competitiveness in two other participants.

> I think this just goes to show that some of you are not doing your fair share and that means the rest of us have to pick up the slack. Diana, creeping up on the number one spot... look out Kate!

About half of these messages do not include a reference to the author of the message at all. Instead, these messages are from spectators who are regular readers of a newsgroup but not necessarily active contributors: they are members of the community, but not necessarily top posters in terms of days active or total posts. These messages often encourage others to try harder or congratulate authors on the event of a large shift in their rank.

Netscan: Tool and Villain

Messages commenting on the Netscan Project comprise 48% of the total sample. Of these, 47% of messages comment on features or bugs in the website, tools, or database. These range from comments about the layout of information on the website and choices in visualizations to insistence that the database contains incorrect information.

In messages from the first several years of Netscan, a surprising number of these posts were coming from rival programmers. Many newsgroups had a participant who served a social role as record keeper for the group. Using a Perl script or even hand tallying message activity, the community score-keeper would report statistics similar to what Netscan now provides instantly. Many of the early feature complaints, then, were from those who once maintained similar programs and those who came to rely on the in-group methods. Chief among the complaints is that Netscan ranks authors based on number of days active, rather than a metric based on posting volume or frequency. Likely hostile over a loss of role in the community, these posters complain about the features and usability of Netscan.

A more common usage in this category is an expression of affect about the tools and their potential, with 54% of messages in this category having an affective quality. These impressions lump into positive, including "cool" or "interesting," and negative, including "Microsoft is evil", bad, scary, and "Big Brother". The positive references outweigh the negative, with 58%, while 28% are concerned about sinister corporate monitoring, 19% see the tools as scary, 10% as bad, and 8% are explicitly hostile to the Microsoft Corporation, with a single message possibly expressing all these subcategories.

Users express a concern for privacy and a fear of how the data might be used. While most users see their Usenet activity as very public in nature, the idea that a company would invest in collecting and analyzing that activity is worrisome to many. Without an idea as to what use the information could be, these individuals assume the worst.

> I have to agree with Liz that this is something of an invasion of privacy too far. Big Brother Bill is watching us....

Although only 7% of messages express this privacy concern, an even smaller population of Usenet participants has chosen to explicitly opt out of the Netscan system. The web interface makes it straightforward and easy to exclude oneself from the publicly available data. Despite this, fewer than 300 have chosen to exclude themselves out of the million or so unique visitors to the website.

4. Discussion

As can be seen in the frequencies of different uses of the Netscan tools, there is a demand for the ability to establish and look up reputations for newsgroups and authors. Here, we discuss what these uses imply and what broader lessons can be learned from this case study.

4.1 Reputation and Community

Several types of newsgroups make frequent use of Netscan as a regular part of their interactions. Top among these are fan and hobby newsgroups. The metrics that Netscan makes accessible allow individuals to achieve several tasks important for the construction of community: boundary maintenance, group comparisons, establishment and enforcement of norms, establishment of regulars and description of their roles, and the identification of threats.

Boundary Maintenance

Common across many senses of community is the need to define and maintain boundaries between one group and another [10]. These practices allow participants to know where they stand relative to social spaces. Entrepreneurs and individuals heavily invested in a particular newsgroup work to define boundaries and bridges around these social spaces to protect them from invasion or to ensure continued connection with other valued spaces.
        Take two cases as examples. In the alt.politics hierarchy, messages are almost never posted to a single newsgroup. Among the most active newsgroups in this category, a typical cross posting

level is 98% of messages being sent between newsgroups devoted to opposing political affiliations. In this space, the boundary is not the individual newsgroup but rather a collection of similarly focused and tightly interconnected newsgroups. The sense of community in this realm is based on vigorous argumentation among a high number of participants. In order to achieve this, threads are deliberately spread across political boundaries in order to incite disagreement and discussion.

On the other side of the spectrum, social support groups in the alt.support hierarchy are highly suspicious and reproachful of cross posting. With several social support groups frequently falling victim to attack from flame warriors, who see a participant in newsgroups like alt.support.autism or alt.support.depression as an easy target for abuse, authors who send a message to several newsgroups simultaneously are distrusted. As these discussions also lean towards more personal conversations, participants are less willing to have what they contribute in what they perceive as a small group spread into potentially hostile "foreign" social spaces.

Group Comparisons

Slightly different from boundary maintenance is the practice of comparing groups to one another. A key distinction is that this practice often comes with a spirit of competitiveness. The most common form of this is found in fan newsgroups for sports teams, though it also appears among newsgroups that share a common topic while remaining distinct newsgroups, thereby competing for participants.

The typical instance of this competitiveness in newsgroup comparisons is a comparison of participation rates between two newsgroups dedicated to different teams. In this way, regardless of how the teams might perform that weekend or that season, the fans feel they have a direct influence on the quality of the franchise. Rather than basing their sense of relative fan commitment on the size of the population attending the games or watching at home, traces of fan activity are evident through Netscan. While fans can do little to rally their team's players, they can directly encourage participation in the related newsgroup.

Norm Establishment and Enforcement

It is important for any community to share a common set of norms
[14]. These allow participants to know how to engage with one an-
other: what is expected, what is encouraged, and what is forbidden.
Since Netscan identifies the most talked about topics in a news-
group, this sense of a community can be gathered at a glance without
a potential new user having to spend a long time reading messages
to get a flavor for the group.

  Many newsgroups cite Netscan data in their periodic posts that
provide descriptions of their purpose, history and regular topics. In
this manner, the author of these posts is reaffirming the standard
practices of the group and highlighting what is both typical and ex-
pected from participants. The ability to view rich information de-
scribing the history of other users also means that new participants
and even long-term participants can be evaluated for their confor-
mity to the norms of a newsgroup. While Netscan does not do the
work of enforcing or establishing norms for a group, the reflective
nature of the data facilitates these tasks for regular participants.

Identifying Regulars and Determining Characters

Another common component of messages that are regularly posted
that describe a newsgroup is the equivalent of a cast of characters.
This behavior varies widely across newsgroups. Many simply paste
the newsgroup's report card directly from the Netscan web site,
which contains a list of top participants ordered by days active.
Others use this list as a basis for lengthy descriptions of each of the
authors. These descriptions explain what one can expect these major
players to do in the group, any quirks they have, and reveal their
personalities to newcomers and fringe participants.

  This practice is a significant resource for the establishment of
community. A newsgroup survives based on its continual accrual of
new messages. There is a huge difference, however, between a
newsgroup which merely receives a large number of messages and
one that receives those posts from a relatively stable core of interact-
ing participants. By listing those participants and describing them,
the report is an implicit encouragement to named participants to re-
main with the newsgroup. At the same time, new participants can

know how the newsgroup participants expect the top posters to be received. This allows the seasoned participants to receive respect without having to continually establish their claim to it by directing less highly ranked posters to show some deference based on their accumulated history as displayed by Netscan.

This cast of characters also establishes the character of the newsgroup. The collection of twenty most prevalent personalities defines the activity and tone of the newsgroup. By identifying both the most active people and the topics they write about, Netscan may be a resource for more stable newsgroup dynamics and smoother introduction of new participants.

Identifying Threats

The potentially anonymous nature of Usenet leaves many users wary of their interactions with others, and correspondingly, Netscan is often used as a resource for performing a kind of background check. This occurs especially when people are considering adopting the query target's advice or entering into a business transaction [11]. The poster typically asks if anyone in the newsgroup is familiar with the other author while simultaneously citing the reputation data available from Netscan.

Another, slightly more common, version of this is in the identification of "flame warriors" and "trolls". If ignored, each of these characters is likely to cease to participate in a newsgroup. If someone takes the bait, however, and responds to the contentious provocations offered by the "troll", the character of a newsgroup can rapidly change. In several newsgroups, an encounter with a "troll" or "flame warrior" can be so distressing that a former regular participant never returns. The reputation data available through Netscan attenuates this possibility by allowing the rapid identification of socially disruptive participants and the encouragement to not engage further with them.

4.2 Identity: The Self and Others

Aside from the effects on the foundation and functioning of community within newsgroups, Netscan is also reflective of individual authors. As such, it helps define an individual's status and role.

Status

As mentioned, a significant use of Netscan is for competitive comparisons of individuals, with 13% of total messages encouraging competition in others and 11% referencing the author with a competitive spirit. This process of making Usenet posts a competitive sport has several important effects.

First, it makes participants more involved in their newsgroups. While originally drawn to a newsgroup for some topical reason, authors often continue to contribute to a newsgroup when a sense of community is present or when participation in the newsgroup fits with the author's sense of identity [2, 3, 5]. The rankings cause many authors to consciously work at increasing their number of days active and their contribution of posts. This raises the level of individual participation and, when the effect is present across several authors, enriches the newsgroup itself. More eyes on the newsgroup translates into shorter waiting times for new messages to receive replies. More posts per author per day yields a more vibrant discussion with more opinions and participants. The shifting rankings from month to month mean that there is an incentive structure that encourages participation over the long-term [27].

The competition for author rank results is also another form of reward for participation. Aside from the sociality of information exchange and dialogue, the sense of "winning" is a separate benefit of participation [23]. This should again raise the level of participation and should reduce rates of exit and turnover.

Reputations

Beyond the competition for rank status among authors, Netscan is reflective of overall reputations. As seen in the data, this allows individuals to defend claims to specific roles within a newsgroup.

This establishment of roles helps to keep individuals in those roles, whether that is as organizer, instigator, or contrarian [5].

The availability of reputation data also helps maintain order within newsgroups [2, 9, 11, 13, 14, 22]. While individual authors can establish their own roles and defend them with long-term data, new participants can also see at a glance how the newsgroup functions as a community. This eases the transition of neophytes into the newsgroup and also likely efficiently deflects individuals who would otherwise have attempted to participate only to find the newsgroup was not a good fit for them.

5. Conclusion

This paper attempts to describe the ways individuals respond to the availability of social accounting meta-data. Using a content analysis of Usenet posts which mention the Netscan reporting system, we find that a large collection of newsgroups use Netscan to monitor the health of their newsgroup, to maintain individual reputations, to investigate the reputations of others, and to facilitate community by lowering costs of entry into a newsgroup and increasing rewards for participation.

While the data express real concern over privacy issues involved in ubiquitous surveillance and monitoring of public spaces like Usenet, the observed behavioral response from users is to use the monitoring to maintain and manage their reputations, rather than attempt to obscure them. The availability of this data has led largely to competitive behaviors, whether between individuals or groups. It has also reduced the level of anonymity experienced as individuals can both share their own reputation and find reputations of others to establish familiarity and a context for interaction.

There are several limitations to this study. Primary among these is the inability to make comparisons to newsgroups which do not mention Netscan or which are entirely unaware of it. Similarly, longitudinal analysis has not been conducted to test for changes in newsgroups over time caused by reputation data. We hope that the largely descriptive work presented here demonstrates the wide range of future research possible in this realm

6. References

1 Brignall T (2002). The New Panopticon: The Internet Viewed as a Structure of Social Control. Theory & Science.

2 Burke PJ, Reitzes DC (1981) The Link between Identity and Role Performance. Social Psychology Quarterly 44 (2): 83-92.

3 Burke PJ, Reitzes DC (1991) An Identity Theory Approach to Commitment. Social Psychology Quarterly 54 (3): 239-251.

4 Burnett G, Bonnici L (2003) Beyond the FAQ: Explicit and implicit norms in Usenet newsgroups. Library & Information Science Research 25:333-351.

5 Cast AD, Stets JE, Burke PJ (1999) Does the self conform to the views of others? *Social Psychology Quarterly* 62:68-82.

6 Coleman JS (1990) Foundations of Social Theory. Cambridge, MA: Harvard University Press.

7 Cook KS (2001) Trust in society. New York: Russell Sage Foundation.

8 Cooley CH (1902) Human Nature and the Social Order. New York: Scribner.

9 Douglas KM, McGarty C (2001) Identifiability and self-presentation: Computer-mediated, communication and intergroup interaction. British Journal of Social Psychology 40:399-416.

10 Erickson K (1966) Wayward Puritans: A Study in the Sociology of Deviance. New York, New York: Wiley.

11 Friedman EJ, Resnick P (2001) The social cost of cheap pseudonyms. Journal of Economics & Management Strategy 10:173-199.

12 Goffman E (1959) Presentation of Self in Everyday Life. Doubleday Anchor.

13 Henderson S, Gilding M (2004) 'I've never clicked this much with anyone in my life': trust and hyperpersonal communication in online friendships. New Media & Society 6:487-506.

14 Hechter M (1987) Principles of Group Solidarity. University of California Press Berkeley

15 Hechter M, Opp KD (2001) Social Norms. University of California Press Berkeley

16 Kinch JW (1963) A Formalized Theory of the Self-Concept. American Journal of Sociology. 68:481-86.

17 Kitts JA (2006) Rival Incentives and Antisocial Norms. American Sociological Review Vol 71 No 2.

18 Kollock P (1999) The production of trust in online markets. Advances in group processes. EJ Lawler, M. Macy, S. Thyne, HA Walker (eds) JAI Press.

19 Lea M, O'Shea T, Fung P, Spears, R (1992) 'Flaming' in computer-mediated communication. Observations, explanations, implications. In M. Lea (Ed.), Contexts of computer-mediated communication. (pp. 89-112). New York: Harvester Wheatsheaf.

20 Mead GH (1934) Mind, Self, and Society. University of Chicago Press.

21 Pfaffenberger B. (2003) A Standing Wave in the Web of our Communications: Usenet and the Socio-Techical Construction of Cyberspace Values. in Lueg C, Fisher D (eds) From Usenet to CoWebs. London, New York: Springer.

22 Resnick P, Zeckhauser R, Friedman R, Kuwabara K. (2000) Reputation Systems. Communications of the ACM. Vol. 43 Issue 12.

23 Resnick P (2002) Beyond Bowling Together: SocioTechnical Capital. Carroll J (ed) HCI in the New Millenium. Addison-Wesley.

24 Rothstein B (2000) Trust, social dilemmas and collective memories. Journal of Theoretical Politics 12:477-501.

25 Turkle S (1995) Life on the screen: identity in the age of the Internet. New York: Simon & Schuster.

26 Turner TC, Smith MA, Fisher D, Welser HT (2005) Picturing Usenet: Mapping computer-mediated collective action. Journal of Computer Mediated Communication. 10(4).

27 Welser HT (2006) A Theory of Status Achievement. Ph.D. Dissertation, Department of Sociology, University of Washington