11870.com: Tagging Site or Social Recommendation System?

Raquel Benbunan-Fich
City University of New York, rbfich@baruch.cuny.edu

Follow this and additional works at: http://aisel.aisnet.org/amcis2009

Recommended Citation
http://aisel.aisnet.org/amcis2009/689

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2009 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
11870.com: Tagging Site or Social Recommendation System?

Raquel Benbunan-Fich
Baruch College, City University of New York (CUNY)
rbfich@baruch.cuny.edu

ABSTRACT
Tagging sites allow users to post information and organize it with tags for future retrieval and to optionally share their entries with others or keep them private. Some tagging sites also allow users to add personal reviews to their entries. Sites that offer publicly available information that includes user-generated reviews are designed to function as social recommendation systems. However, since sharing is voluntary and reviews are discretionary and laborious to produce, it is not clear whether a site with these optional features can function as a social recommendation system. Using activity data from a site with these characteristics (11870.com), we test whether contributors use it as a social recommendation system. We find that the prevalent user profile is that of a public contributor, for whom the proportion of entries annotated with reviews is 68%. Almost 40% of the public contributor base, particularly newer users of the site, provides reviews for all of their entries. Our results indicate that this tagging site is increasingly viewed as a social recommendation system despite the discretional nature of sharing resources and adding reviews. An important implication of these findings is that optional features do not undermine the ability of a tagging site to function as a social recommendation system.

Keywords
Social Recommendation Systems, Tagging, User-generated content.

INTRODUCTION
One of the main characteristics of Web 2.0 sites is that they enable users to upload their own content and share it with others. The type of content varies widely and can include bookmarks (Benbunan-Fich and Koufaris, 2008), product reviews (Peddibohla and Subramani, 2007), encyclopedia entries (Nov, 2007), photos (Marlow, Naaman, Boyd and Davis, 2007), and videos. Some Web 2.0 sites are designed to function as recommendation systems. The recommendation function relies on user-generated reviews of products or services such as books, movies, restaurants and hotels. In addition to relying on user-generated content, a common feature of the new generation of Web 2.0 sites is the availability of tools to organize and filter user-generated content with freely chosen descriptive words or tags. In most sites, users can only tag the content they post, while in a few other sites such as Flickr, for example, users can tag their own photos as well as those posted by others.

Through tagging, users organize their contributions and make them available to others. Most tagging systems offer three sharing options: (1) public - to make their contributions available to all other users of the site, (2) semi-private - to make their contributions available to their designated friends or contacts only, and (3) private - to keep posted information for their personal use only. Social tagging systems are created when contributors share their entries with the public of the site. Users who make their tagged resources available to others help create and maintain an online public repository of catalogued entries. When people search this repository using tags, they can find their own entries along with those stored by others under the same tags (Golder and Huberman, 2006). This serendipity effect leads users to discover or find unexpected co-located resources sharing the same tag, and it is a significant collective benefit of social tagging systems (Riddle 2005, Weinberger 2007).

The nature of the resources stored in a tagging system determines, to a large extent, the function that the site fulfills. For example, if the site provides tagged bookmarks, as in del.icio.us, its main objective is to bring websites to the attention of others (Benbunan-Fich and Koufaris, 2008). If the site offers tagged photographs, as in Flickr, its main objective is to allow sharing of digital pictures among users. Occasionally, the object of the site is more subjective and instead of digital objects, users post and share opinions. For example, in ePinions users can share their reviews of consumer products and in RealEats, patrons can post reviews of restaurants. In these cases, the main function of the site is to provide social recommendations with user-generated entries that include basic descriptive information of the product or service, classification tags and customer reviews. In particular, review sites of restaurants, hotels, attractions and other services require users to post and tag their entries, add reviews and make their entries available to others. The absence of tags would make public entries difficult
to find, and the absence of reviews would make the site closer to a web-based user-generated directory of products and services, than to a social recommendation system.

This study is focused on tagging sites where: (1) Public contributions are optional and individuals can use the site exclusively for themselves (i.e. privately), or to share resources either with a group of known contacts (semi-private) or with all other users of the site (public). (2) User-generated content includes at the very least descriptive information about a physical product, place or service. This information can be optionally tagged and accompanied by a user-generated review (i.e. tags and reviews are not mandatory). The research question guiding this study is whether a tagging site that offers different sharing options (private, semi-private and public) and where posting reviews is voluntary can function as a social recommendation system. We articulate three conditions for a site to function as a social recommendation system and examine them with activity data of actual users of a tagging site (11870.com). The next section provides the theoretical background and outlines the three conditions for a site to function as a recommendation system. The following section describes the research methodology. The data analysis and discussion of results are presented next. The paper concludes with implications, contributions and future research directions.

THEORETICAL BACKGROUND

In recent years, user-oriented online sharing communities have experienced exponential growth. Some of these communities are based on participant interaction and reciprocal exchange of information. Typical examples include bulletin boards, discussion groups, help forums, and open source groups. Even though most of the participants do not know each other, users are able to directly interact with others through the system by answering their requests for information or help. This direct interaction between users can partially explain why participants offer voluntary contributions to benefit other users of these communities. These acts are evidence of pro-social behavior in an interpersonal environment.

In contrast, other types of online sharing communities are more impersonal because each user interacts with the system as a whole but not with specific other users. Individual contributions are voluntary and do not respond to any particular request for information. Examples of these sites include social bookmarking systems (such as del.icio.us) and product reviews (such as Epinions.com), where individuals make contributions to a shared repository of information, often without direct user interaction. Despite the lack of direct personal interaction, empirical studies have shown that people exhibit pro-social behaviors in these sites and that their contributions are aimed at helping other users. For example, in Amazon.com, users’ contributions of product reviews are voluntary and intended to benefit other users instead of the contributor, who is already in possession of the information (Peddibohtla and Subramani, 2007).

This study is concerned with repository-driven tagging systems. Specifically, we examine a type of Web 2.0 site based on user-generated content organized with tags. Tags are the means for organizing and indexing contributions and connecting people with resources. Users can classify a particular resource – such as a web page, a blog post, an image, a reference to a physical product or service, or any other type of object – with their own keywords or tags. Through tagging, users share their resources with others. Users are connected to the resources they post because each contribution is identified with the username of the contributor. Resources sharing the same tags are co-located in the repository. Thus, tagging systems introduce a new modality of indirect social communication as they connect people through the tagged resources they make available to others (Marlow et al., 2006).

According to Marlow et al. (2006), the design characteristics of tagging sites influence the type of content users are willing to contribute and the resulting information dynamics. A tagging site with different sharing alternatives (public, private and semi-private) and optional additions of tags and reviews to the basic content lends itself to a variety of purposes. When users post basic entries (i.e. only object descriptors), the site is used for web-based storage of information. If basic entries are further described with tags, users are ensuring future access and easy retrieval of their original entries. Depending on the selected sharing option, these entries may be intended for personal use, for small group use or for public use. When users add their own opinions or reviews to their entries, the site is used as an annotated personal organizer (if private), as a small-group recommendation system (if semi-private) or as a social recommendation system (if public).

The research question guiding this study is whether tagging sites offering the public-private choice for posting and where reviews are optional can become social recommendation systems. The basic premise underlying this question is that users are rational actors for whom the expected benefits associated with their contributions must outweigh production costs (Kankanahalli, Tan and Wei, 2005). Costly contributions in terms of time and effort must be associated with greater expected benefits than relatively simple contributions. For example, contributions that require users to describe an object (or a place), tag it and write a review are more costly to produce than those that only require a brief entry. Therefore, users have to expect benefits that outweigh the costs of contributing if they provide reviews along with their entries.
Private Users vs. Public Contributors

Kollock (1999) provides two alternative views of online contributors depending on their motivations. At one extreme, users are viewed as egoistic individuals exclusively concerned with producing resources for their own benefit and later consumption. At the other extreme, users are assumed to be altruistic and concerned about addressing the needs of others with the resources they produce. These assumptions result in two different views of users of tagging sites. Selfish individuals will add information to a tagging site for their own personal benefit, and will use the site to fulfill his/her own personal needs of keeping track of resources. In contrast, selfless individuals will contribute to the site to help others and share information. Accordingly, selfish individuals will tend to keep the resource they contribute for their personal use (private), while selfless contributors will use the site mainly to share resources with others (public). This dual purpose is consistent with the organizational and social taxonomy for using tagging sites articulated by Marlow et al. (2006). Private users view the site as an organizational tool for storing, cataloging and documenting resources for their personal use. In contrast, public contributors use the system for social purposes by making their resources available to others. Although public contributors also organize their resources, their main objective is sharing.

Both types of producers (public and private) are rational actors that will try to minimize costs and maximize benefits associated with the production of resources for the site. Private users differ from public contributors in the benefits they expect from using an online tagging system. Private users derive benefits from the organizational features of the site (storage and future retrieval), while public contributors obtain rewards from revealing their contributions and sharing them with others. Thus, individuals decide on their behavior based on expectations of benefits associated with their use of the system. Social exchange theory explains the rationale for sharing information with others (Blau, 1964). People treat information sharing like other exchanges because there is an expectation of some future return from sharing. Such returns include reciprocity expectations, better reputation, and/or increased power (Wasko and Faraj, 2005). In addition, information sharing is also influenced by the social context where it occurs (Constant et al., 1994). People may be inclined to contribute due to concerns for the needs of others (Kollock, 1999), for the well-being of the community (Jarvenpaa and Staples, 2000) or for other altruistic reasons (Parameswaran and Whinston, 2007).

Production costs are independent from the private or public user profile because they are a function of the type of resource produced (i.e. object descriptor, tag, or review). Entering object descriptors is necessary to use the site. However, such descriptors typically involve typing a short name or copying/pasting a link and therefore involve minimum effort. Similarly tagging, which is necessary to index and find the resources stored at the site, is relatively easy because it is based on the human ability to naturally classify objects with labels. In contrast, writing a review is a more costly endeavor, as it requires producers to express their opinions or describe their experiences in writing and in a coherent way.

Comparatively, tagging and reviewing an entry require different levels of effort and hence have different production costs. With tags users perform an indexing function by allowing resources to be classified and retrieved at a later time, while reviews annotate entries that include the name and address of the place, and customer reviews. In order to post an entry, users need to sign up and create their own account. Each entry can be designated by the contributor as private, semi-private (to share only with friends – called contacts), or public (to share with all other users). Depending on the sharing option chosen by

---

1 The name of this site comes from the telephone information service number in Spain (118). Although it is a completely web-based service, when the owners secured the telecommunications license to provide information on the Internet, they were assigned this number and decided to use it as the name for their site.
users, the site can serve as a private keepsake of places (such as hotels, restaurants, museums, etc.) and services (such as painters, electricians, plumbers, etc.). As such, it would function as a personal/virtual Rolodex. Alternatively, the site can be used as a social recommendation system when user-generated reviews are shared with the public at large. Although these two functions – personal virtual Rolodex and a social recommendation system – can coexist, we are interested in the predominant pattern of use of the site.

Of particular interest is the analysis of contribution activity as well as the use of the tagging and reviewing features, as these features will indicate the patterns of use of the site. For this study, the owners of the site shared their entire user activity database with us, masking the names of their users to protect their anonymity. Activity data included the following information: number of public, private or semi-private entries, number of tags, number of reviews, joining date, last login date and last post date. With this information, we calculated the following variables: total entries, percentage of entries made public, average tags per entry, proportion of entries with reviews, and tenure at the site in days (data collection date minus join date). These variables are measured at the user level.

**ANALYSES AND RESULTS**

On the day of the data collection, the user database contained a total of 6,264 users from which 2,260 had an account but had not made any contribution to the site. After removing zero producers, the sample consists of 4,004 users. Collectively, these users contributed a total of over 50,000 entries (places or services) from which almost 45,000 are public, approximately 63,000 tags and nearly 28,000 reviews. The percentage of public contribution is, on average 91%, indicating that most users contribute all their entries to the public repository. The number of tags is on average almost 16, with a large standard deviation of 47.89, indicating a wide range of tagging efforts. In terms of length of membership at the site (tenure), the average is 230 days from the date the user joined until the day of the data collection. The average last posting date is 160 days and the average last login date is 131 days.

A total of 3,333 (83%) producers are pure public contributors with 100% of their entries publicly available at the site. At the other extreme, 148 members are pure private users, who keep all of their contributions to themselves. A very small percentage (2%, 67 users) are pure semi-private, who share entries only with their friends. The remaining 456 (11%) individuals are mixed users who exercise different sharing options as they make contributions to the site. Using a chi-square test for equal proportions, we find that the difference among the percentage of users in each profile is significant ($\chi^2 =7327.89, p<0.0001$). Thus, the prevalent user profile is that of the pure public contributor followed by mixed users. Further analyses of mixed users reveal that an average mixed user designates about 66% of his posts as public, keeps 23% of his entries private and selects 11% for sharing with his contacts (i.e. semi-private). This average profile suggests that mixed users have a preference for public posts over private or semi-private contributions.

Although 83% of the producer segment consists of pure public contributors, there are noticeable differences in the amount of public contributions they make. The range of public contributions goes from 1 to 615 entries, with an average of 11 and a median of 2. This disparity in contribution levels is consistent with prior studies of user-generated content (Peddibothla and Subramani, 2007). In 11870.com, a few top ranked users make very large contributions of public resources and there is a long tail of small contributors, most of which have contributed only one entry upon joining the site and have not returned since. For the rest of our analyses, we excluded these unitary contributors and focus on the remaining group of pure public contributors, who have posted more than one entry to the site.

In the group of 1,874 pure public contributors with more than one entry, the average number of tags per entry is 1.78. Thus, on average, contributors tend to index their entries with almost two tags. In contrast, the percentage of entries annotated with reviews is less than one, as contributors do not provide a review for each one of their posts. For the average public contributor, the proportion of entries annotated with reviews is 68%.

Based on their average tags per entry, we divide this group of 1,874 contributors into heavy-taggers and light-taggers. Heavy-taggers are those with an average of one or more tags per entry, and light-taggers are those with less than one tag per entry. A chi-square analysis shows that heavy-taggers significantly outnumber light-taggers (67% vs. 33%; $\chi^2=211.79, p<0.0001$). A similar analysis is performed with the level of reviewing and the sample is divided into heavy-reviewers and light-reviewers. In this case, heavy-reviewers are those who enter a review for each one of their entries, while light-reviewers are those who do not review all the entries they post in the system. In contrast to the analysis for tagging, a chi-square test of reviewing activity reveals that light-reviewers significantly outnumber heavy-reviewers (58% vs. 42%; $\chi^2=45.498, p<0.0001$). Nonetheless, it is worth noting that 42% of contributors post reviews for all their entries.

A cross tabulation of the number of contributors in each category shows that the lowest amount of users (10%) is found in the light-tagger/heavy-reviewer category, while the largest amount (654 contributors or about 35%) is found in the heavy-
tagger/light-reviewer cell. The second most populated cell is the heavy-taggers/heavy-reviewers category with 598 contributors (32%). A chi-square test indicates that these proportions are significantly different ($\chi^2=47.71$, $p<0.0001$). This analysis indicates that most contributors predominantly tag their entries and add occasional reviews (heavy-tagger/light reviewer), or engage in both tagging and reviewing for all their entries (heavy-tagger/heavy-reviewer). These results are presented in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>L-Taggers</th>
<th>H-Taggers</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-Reviewers</td>
<td>429 (22.9%)</td>
<td>654 (34.9%)</td>
</tr>
<tr>
<td>H-Reviewers</td>
<td>193 (10.3%)</td>
<td>598 (31.9%)</td>
</tr>
</tbody>
</table>

$\chi^2=47.71$, $p<.0001$

Table 1. Cross-Tabulation of Tagging vs. Reviewing

To gain further insights into the production patterns of users in terms of their tagging and reviewing activity, we conducted one-way analyses of variance. These tests appear in Table 2. We find that contributors with an average of one or more tags per entry (Heavy-Taggers) have posted fewer public contributions than Light-Taggers (12.64 vs. 18.07) but have been members of the site for a longer period time (261.79 days vs. 230.44 days). Light-Taggers, however, are more frequent contributors than Heavy-Taggers as they show significantly lower lag time between their last posting and the data collection date (141.70 vs. 156.28 days). In terms of reviewing activity, occasional reviewers (Light-Reviewers) are logging into the system more frequently and posting resources more often than those who post reviews for all their entries (Heavy-Reviewers). Although Heavy-Reviewers have fewer public contributions that their counterparts, they have joined the site more recently.

<table>
<thead>
<tr>
<th></th>
<th>L-Taggers (n=622)</th>
<th>H-Taggers (n=1,252)</th>
<th>$R^2$</th>
<th>Model F (p-level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Contributions</td>
<td>18.07</td>
<td>12.64</td>
<td>0.6%</td>
<td>11.02** (p&lt;.0009)</td>
</tr>
<tr>
<td>Days since Joining</td>
<td>230.44</td>
<td>261.79</td>
<td>1.6%</td>
<td>30.43***</td>
</tr>
<tr>
<td>Days since Last Login</td>
<td>111.24</td>
<td>118.30</td>
<td>Model not significant</td>
<td></td>
</tr>
<tr>
<td>Days since Last Posting</td>
<td>141.70</td>
<td>156.28</td>
<td>0.03%</td>
<td>6.01*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>L-Reviewers (n=1,083)</th>
<th>H-Reviewers (n=791)</th>
<th>$R^2$</th>
<th>Model F (p-level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Contributions</td>
<td>20.61</td>
<td>5.99</td>
<td>4.7%</td>
<td>91.65*** (p&lt;.0001)</td>
</tr>
<tr>
<td>Days since Joining</td>
<td>257.68</td>
<td>242.78</td>
<td>0.4%</td>
<td>7.47** (p=0.006)</td>
</tr>
<tr>
<td>Days since Last Login</td>
<td>103.25</td>
<td>133.26</td>
<td>1.8%</td>
<td>33.72***</td>
</tr>
<tr>
<td>Days since Last Posting</td>
<td>137.86</td>
<td>170.04</td>
<td>1.7%</td>
<td>32.64***</td>
</tr>
</tbody>
</table>

Significance levels: *** $p<.0001$; ** $p<.01$; * $p<.05$

Table 2. One-way Analysis of Variance

Overall, the analyses presented above show that the majority of users of 11870.com fit the pure public contributor profile and that these contributors tag and add reviews to their entries with a preference for tagging over reviewing. Almost a third of the contributor base consists of users whose contributions take more time and effort, such as Heavy-Taggers and Heavy-Reviewers. In addition, newer users add reviews for all their entries.

---

2 Two-way analyses of variance on the same factors show no significant interaction effects.
DISCUSSION

In this study, we investigate whether a tagging site with three choices for sharing contributions (private, semi-private and public) and where posting reviews is optional, functions as a social recommendation system or as a web-based organizer of tagged information. Based on our discussion of two different user profiles (public contributors vs. private individuals) and cost-benefit considerations, we proposed three conditions for the site to function as a social recommendation system, namely availability of public resources, indexed with tags and accompanied by reviews. We analyzed activity data of a tagging site (11870.com) with these features and found that the private and semi-private sharing options do not undermine the public repository as the majority of users make most of their entries available to all other users.

Although we identified different user profiles (pure public, pure private, pure semi-private and mixed), the proportion of public contributors is significantly higher than the proportion of other types of users. Moreover, even those who selectively share their contributions with all others (mixed users) exhibit a preference for public posts over private or semi-private entries. Taken together these findings suggest that most individuals use the site to share their resources with others. Furthermore, in this site, a small group of top contributors are responsible for the majority of public contributions of places and services, and there is a long tail of users that make marginal or no public contributions at all. This result indicates that the abundance of resources in this site is produced by a small group of very prolific contributors, as it is the case for other online repositories based on user-generated content (Peddibohla and Subramani, 2007).

Given the abundance of public resources resulting from the predominant public contributor profile and the prolific efforts of a small group of top contributors, we examined if those resources were cataloged with tags and accompanied by reviews. We found that the average user of 11870.com tags his public contributions with almost two tags and adds reviews for 68% of his entries. We also found that users whose contributions take more time and effort to produce – because they are annotated with reviews – tend to contribute less than their counterparts who do not include reviews with all their entries. This finding is consistent with the assumption of cost minimization. A more detailed comparative analysis of tagging and reviewing activity shows that almost a third of the contributor base consists of users whose contributions take more time and effort, such as Heavy-Taggers (those who add one or more tags per entry) and Heavy-Reviewers (those who enter reviews for all their entries). Overall, about 40% of the user base provides reviews along with their entries. Remarkably, members with longer tenure at the site tend to add more tags, while newer members tend to post reviews for all their entries. Taken together, these findings indicate that while the site may have been used initially for tagging, it is increasingly used to provide recommendations. These are encouraging signs for the success of the site as a social recommendation system.

The implications of these results are manifold. First, when online systems are designed with optional sharing features and optional reviews, the patterns of use of the site will depend upon the predominant user profile. The social function of the site is fulfilled when there is an abundance of public resources because contributors outnumber or over-produce resources for sharing. Therefore, website owners who want to transform their sites into online spaces for information exchange should encourage and reward public contributions by featuring prolific and/or frequent public contributors in the site’s main page or by offering rewards associated with the contributor status. Second, the organizational function of the site is the foundation for sharing. Tagging systems should be designed to ensure that entries are tagged consistently and meaningfully while reducing tag production efforts in the process. This can be achieved through tag suggestions upon entering resources or with default tagging for new entries. Third, organization and sharing do not result in social recommendation systems unless users are encouraged to enter reviews and rewarded for doing so. One alternative to promote the addition of reviews is to offer digital awards (medals, stars, points, etc.) for thorough reviewers and/or for most useful reviews.

Our results should be interpreted in light of the limitations of the study. First, the analyses are performed with objective activity data kept at 11870.com and therefore lack information about the demographic characteristics or psychological motivations that could explain the drivers of different types and levels of user contributions. Further research should explore the motivations of users who voluntarily contribute content to Web 2.0 sites and investigate in more depth the perceived cost and benefits associated with different types of contributions. Second, our data comes from users of only one tagging site and may not be replicable in other contexts where the nature of contribution is different or where there are different sharing options. Third, our examination is carried out from the perspective of the producer of resources, as opposed to the consumer of such resources. Accordingly, a key assumption is that if the system provides user-generated reviews, it will function as a social recommendation system. However, studies of actual usage and usefulness of the reviews are likely to ultimately confirm whether these systems are truly used as social recommendation systems. Despite these limitations, this study advances our understanding of versatile Web 2.0 sites that could function as private tagging systems or as social recommendation systems and paves the way for further research.
CONCLUSION

Social tagging systems are one of the many popular applications of the current generation of Web 2.0 sites that give users the tools to write, post, tag and upload their own content turning them into contributors instead of passive consumers of information. Some social tagging systems offer the option to post and tag resources privately, semi-privately or publicly. Given the possibility of pure private or semi-private behavior by users of these systems, and the optional nature of providing reviews with their entries, we examined whether such sites function as simple repositories of tagged resources or as a social recommendation systems. We found that increased system versatility in the form of public-private choice for posting and optional reviews do not undermine the ability of the site to function as a social recommendation system. Our results indicate the web is voluntarily used a medium for sharing information and providing advice.

REFERENCES