Harnessing Wisdom of the Crowds Dynamics for Time-Dependent Reputation and Ranking

Elizabeth M. Daly
IBM, Dublin Software Lab
IBM Software Group
Elizabeth_Daly@ie.ibm.com

***Abstract*—The "wisdom of the crowds" is a concept used to describe the utility of harnessing group behaviour, where user opinion evolves over time and the opinion of the masses collectively demonstrates wisdom. Web 2.0 is a new medium where users are not just consumers, but are also contributors. By contributing content to the system, users become part of the network and relationships between users and content can be derived. Example applications are collaborative bookmarking networks such as del.icio.us and file sharing applications such as YouTube and Flickr. These networks rely on user contributed content, described and classified using tags. The wealth of user generated content can be hard to navigate and search due to difficulties in comparing documents with similar tags and the application of traditional information retrieval scoring techniques are limited. Evaluating the time evolving interests of users may be used to derive quality of content. In this paper, we propose a technique to rank documents based on reputation. The reputation is a combination of the number of bookmarkers, the reputation of the bookmarking user and the time dynamics of the document. Experimental results and analysis are presented on a large collaborative IBM bookmarking network called Dogear.**

***Index Terms*—Reputation, Ranking, Social bookmarking, Social Search**

I. Introduction

Web 2.0 is a new medium where users are not just consumers, but are also contributors. By contributing content to applications, such as Flickr or del.ic.ious, users become part of the network and relationships between users and content can be derived. These social networks between people and content can potentially be harnessed to capture "the wisdom of the crowds" [10]. Example applications are collaborative bookmarking networks such as del.icio.us[1] and file sharing applications such as YouTube[2] and Flickr[3]. These networks rely on content being classified by user generated tags and thus provide a mechanism for informal catagorisation. The rating and ranking of documents in such networks is problematic. Search engines determine rank based on indexable content, in a tagging environment indexable content may be scarce. Google's page rank uses the connectivity of the entire network to influence rank. However, in a tagging based environment content may be an isolated piece of data, with little or no connectivity network between data sources. Additionally, the wealth of user generated content can be hard to navigate and search due to difficulties in comparing documents with similar tags.

One solution is to take into account the popularity of a document. A document that has been consumed by a large number of users can be deemed popular where a 'boost count' denotes of the number of consuming users. However, this technique inevitably favours older documents. As a result, newly added content cannot compete with documents consumed by a large number of users in the past. Consequently, ranking based on popularity can lead to a 'rich-get-richer' scenario suppressing newly added documents. Cho and Roy demonstrated that ranking based on popularity hinders the discovery of new web pages and that increases in web page popularity are heavily influenced by search engine ranking [3]. To overcome this bias towards older documents, some applications allow ranking based on recency, where content is ranked based on how recently the content was added to the network. However, ranking based on recency neglects the quality of a document and makes discovery of relevant and reputable content difficult.

In this paper, we propose a technique to rank documents based on reputation, illustrated in figure 1. In order to harness the "wisdom of the crowds" the reputation of users and the reputation of documents are integrated into a single metric. Consequently, reputable users, e.g., trend-setters have a greater influence on the reputation of bookmarked documents than users with a low reputation. To address the age bias, both reputation metrics undergo a decay process. This means that inactive users and documents decrease in reputation over time with the advantage of capturing the current trends. Experimental results and analysis are presented on a collaborative IBM bookmarking network called Dogear.

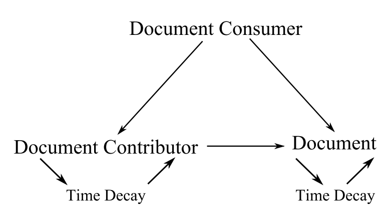


Figure 1: Reputation Network

[1] www.delicious.com
[2] www.youtube.com
[3] www.flickr.com

978-0-7695-3689-7/09 $25.00 © 2009 IEEE
DOI 10.1109/ASONAM.2009.69

II. Social Ranking Based on Reputation

A Reputation Ranking is proposed which captures the time dynamic reputation of a bookmarked document. This work is inspired by a Delay-Tolerant Network (DTN) routing algorithm which measures the predictability of mobile node encounters [8]. Four factors are integrated into the reputation of a document:

1) The number of users consuming the document;
2) The reputation of the user contributing the document;
3) The time dynamics of user consumption;
4) The time dynamics of consumption of documents contributed by the user.

Incorporating these factors into a ranking system and coupling the popularity of documents along with the timeliness of recent bookmarking activities shifts the search results from mainly dominated by old results to more recent up-to-date results. The advantage of such an approach manifests itself in the discovery of trends, which otherwise is more difficult to achieve.

A. User Reputation

Users consume content contributed to the network by other users. The number of users consuming a given contributors content may be seen as a form of implicit recommendation. If a user consistently adds content that other users deem interesting, then this user can be considered a reputable contributor. The reputation of a user is initialised to an input parameter $R_{init} \in [0,1]$. After this time, the value is updated with a reward constant R_{reward}

$$R_{new} = R_{old} + (1 - R_{old}) \times R_{reward}. \tag{1}$$

Equation 1 is applied to a user every time a person adds a document to their collection contributed by that user. The selection of the reward constant R_{reward} should consider the consumer rate of the application. If the number of expected consumers is in the order of hundreds or thousands, then an overly high value of R_{reward} will potentially cause popular content to quickly converge towards 1 making it difficult to differentiate between similarly popular content.

B. Document Reputation

The number of users that add a document to their collection indicates popularity and may be used to derive the quality of the document. A simple reward model is used to measure a document reputation. Each time a user adds a bookmark to their library, the reputation of the document is updated, reinforcing the value of the bookmark. As such, a bookmark's reputation can be measured in a similar manner to how user's reputation is measured by applying equation 1.

C. Time Dynamics

Golder and Huberman researched bookmark popularity and reported that 67% of pages reached their peak popularity levels in the first 10 days after being added to del.icio.us [4]. However, 17% took over 6 months in order to reach peak popularity.

In order to capture the time dynamics and current relevance of the document to the user population, the reputation value is decayed over time. Therefore, documents that are continuously being added to user libraries are rewarded, and inactive content is degraded by:

$$R_{new} = R_{old} \times \gamma^{k}, \tag{2}$$

where γ is the decay coefficient and k is the number of elapsed time units since reputation value was last aged. As a consequence, bookmarks with a high reputation are decayed over time unless user consumption remains steady. The selection of the decay coefficient, γ, and time unit, k, should be based on the importance of recency in the application. An aggressive decay model can be used to detect short term trends with a low value of γ and a time unit of days or even hours. An application where recency is less important than the quality of content, a more conservative decay model may be employed with a value of γ tending towards 1 and a time unit of weeks or months. As with a document reputation, a user's reputation must be time dependent in order to reflect the recent nature of their contributions. As such, user reputations are decayed over time using equation 2 also.

D. Reputation Ranking: combining document and user reputation

As shown by Golder and Huberman, documents can differ greatly in the rate of user consumption. Evaluating the quality of newly contributed content is problematic when relying on a simple boost count. As a consequence, the proposed solution of Reputation Ranking given in equation 3 utilises the reputation of a document in conjunction with the reputation of the user consuming the document. The Reputation Ranking of a bookmark document is denoted as:

$$R_{new} = R_{old} \times R_{bookmarker} \times \beta, \tag{3}$$

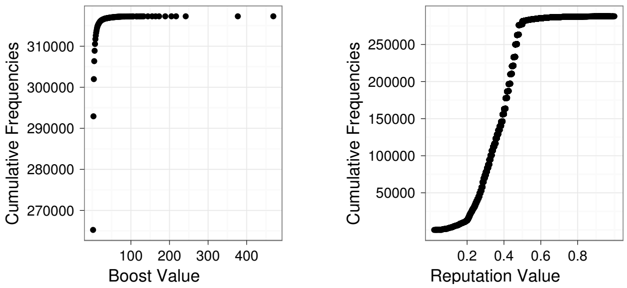
where β is a weighting constant representing the extent to which a consuming bookmarker's reputation may influence a bookmark's reputation. The value of β determines the amount of influence a user's reputation may have over a document's reputation value.

III. Experiment Results

In this section the experiment used to evaluate Reputation Ranking is described and the performance is compared to basic boost ranking.

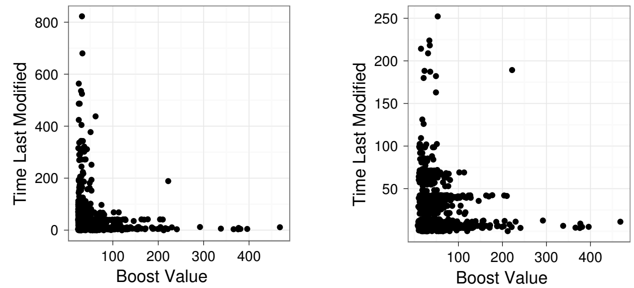
Table I: Dogear dataset

Number of Users	Number of Bookmarks	Number of URLs
10259	505472	317362



(a) Boost Ranking (b) Reputation Ranking

Figure 2: Cumulative distribution of ranked collection



(a) Boost Ranking (b) Reputation Ranking

Figure 3: Ranking of the top 1000 documents

A. Experimental Setup

In order to evaluate the premise of Reputation Ranking a large collaborative bookmarking data set is used. IBM's collaborative bookmarking solution Dogear [9] is popularly used by IBM employees. The Dogear data set contains an extensive network of users and contributed URLs, shown in table I. The time dependent additions of bookmarks to the network are used to calculate the bookmark and bookmarker's reputation. This is achieved by simply replaying the bookmarking and tagging activities in chronological order.

Figure 2 shows the cumulative distribution of ranking based on boost of the entire document collection. As can be seen the graph obeys the power law with a power law coefficient of 2.46 which results in a very long tail with a small number of highly ranked documents.

Upon inspection 99% of the documents have been bookmarked by 13 people or less. As a result, the majority of documents are differentiated by only a small number of additional users bookmarking the document and therefore are difficult to compare. The rank distribution using reputation is cumulative normal which results in a clear divide between highly reputable documents and documents with low reputation.

In order to evaluate the benefits of Reputation Ranking the result set is examined compared to boost ranking. The following details are evaluated:

- Boost count: This is a good representation of the popularity of a document, however, evaluating based on that alone rewards a rich-get-richer ranking algorithm and needs to be viewed in the context of document recency also.
- Time since last modified: A document with high boost count, though popular, may not be highly relevant if a long period of time has passed since the document has been active in the network, i.e., the document has been bookmarked.
- Document life-span: This is the difference between the time the document has last been active compared to when a document was first added to the network. Boost tends to favour documents that have been in the network for long periods of time, Reputation Ranking on the other hand aims to reduce this life-span dependency.
- Distribution of person reputation: In order to reduce the life-span dependency, other measures need to be taken into account in order to determine the quality of a newly contributed document. This is achieved through the additional reward of documents bookmarked by users with a high reputation.

Table II: Reputation Rank parameter values

R_{init}	R_{reward}	γ	k	β
0.5	0.1	0.98	number of months	0.2

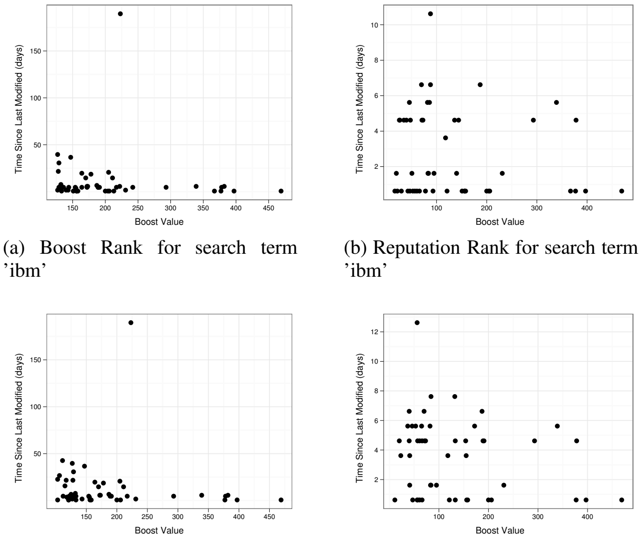
Table II shows the parameter values used to calculate the reputation of documents and users. Each document and user is initialised to a value of 0.5, i.e., no user or document is considered to have a different ranking until either user uptake or time determine otherwise. A relatively low decay model was chosen where the time unit of decay is months.

B. Document Ranking

In order to demonstrate the time dependent benefits of reputation ranking figure 3 shows the time since the document was last added to the network, compared to ranking based on boost value of the top 1000 documents in the network. Boost ranking returns documents that have not been updated in 2 years. Reputation ranking returns significantly more recent results, while still returning documents with a high boost count that have been recently updated. The overlap of results returned is 63%. Older results with a high boost value, have been replaced by results with a lower boost value but that have been added recently.

C. Ranking Popular Documents

In order to demonstrate ranking highly popular documents, the most frequently used tags across the document collection were selected that return the most results. Figure 4 shows the top 50 results returned based on a tag search for the two most popular tags, in this case 'ibm' and 'web2.0'. The trend clearly shows that ranking based on boost returns a



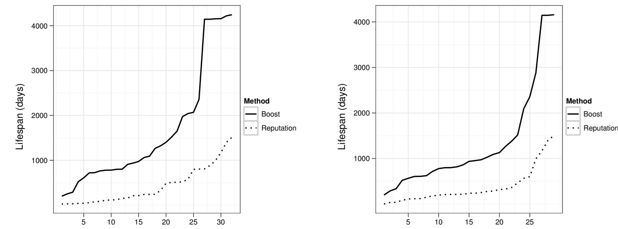
(a) Boost Rank for search term 'ibm'

(b) Reputation Rank for search term 'ibm'

(c) Boost rank for search term 'web2.0'

(d) Reputation rank for search term 'web2.0'

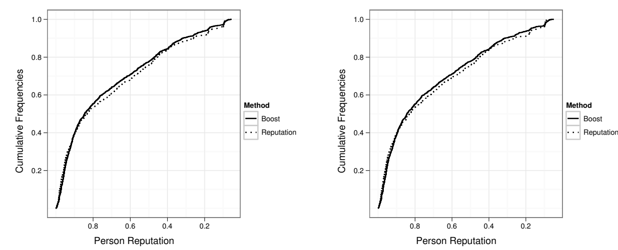
Figure 4: Top 50 ranked document



(a) search term 'ibm'

(b) search term 'web2.0'

Figure 5: Age of documents not overlapping in results



(a) search term 'ibm'

(b) search term 'web2.0'

Figure 6: Person Reputation distribution in non overlapping results

number of older bookmarks when compared to Reputation Ranking. Upon analysis, the search results returned by the ranking schemes have an overlap of 38% and 44% for 'ibm' and 'web2.0' respectively. The overlap is low, because the maximum time since a document was last modified is 11 days for Reputation Ranking and as high as 190 days in boost ranking. This highlights the fact that Reputation Ranking includes more than 50% recent documents that were ignored by boost count ranking.

As previously discussed, ranking based on boost count significantly favours older documents with a long life-span. In order to evaluate this observation, the results of both ranking mechanisms are examined and only the non-overlapping documents are analysed. Figure 5 shows the life-span of the non-overlapping results of both Reputation Ranking and boost ranking. In both cases, boost rank has returned documents with a high life-span, whereas Reputation Ranking returns documents with a shorter life span but that have been updated more recently.

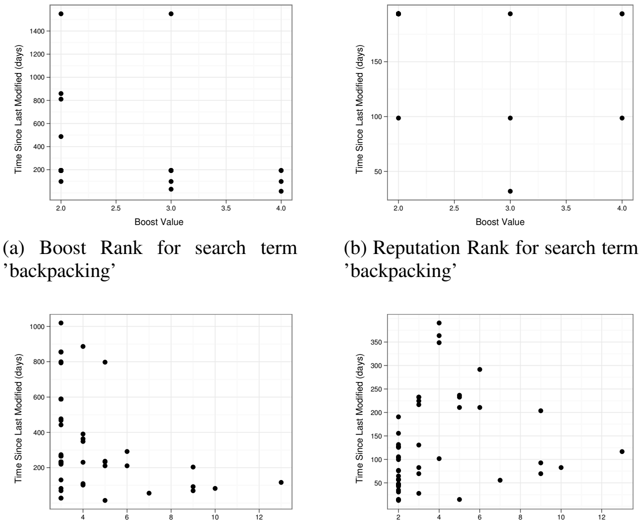
D. Ranking Less Popular Documents

As was seen in figure 2, 99% of the documents in the network are bookmarked by 13 people or less. As a result ranking based on a simple boost count has limited meaning as there is no clear divide between the document ranks. If 200 documents have been bookmarked 3 times, the selection of the top 50 results is arbitrary. In order to demonstrate the ranking behaviour, the result set is limited to documents that have been bookmarked 13 times or less. This is achieved by selecting the two most frequently used tags that occur exclusively in documents that have been bookmarked 13 times or less, where more than one person has applied the tag [4]. In this case the two terms are 'backpacking' and 'photoshop'. The disadvantage of ranking based on boost count is most obvious when there is little to differentiate results once the boost count drops to 2. Additionally, some results are as old as 4 years which are clearly not current and may no longer be relevant. Reputation Ranking on the other hand strikes a balance between recent additions and frequently consumed documents.

Overlap for the search term 'backpacking' is 82% and so there are only a few outliers that are included in the non-overlapping result set shown in figure 8. The age of the documents returned by Reputation Ranking are all zero, as such, the reputation of the user consuming the document has been a deciding factor in ranking these results. The overlap for the search term 'photoshop' is 46%.

Figure 9 shows the distribution of the reputation of users that bookmarked the documents in the non-overlapping results set. In figure 9 a) the reputation of the contributing users

[4] We discarded two tags that were used by a single author to tag all content in their collection. The results showed identical behaviour, however we felt the utility of searching based on a personal tag was limited to that user.



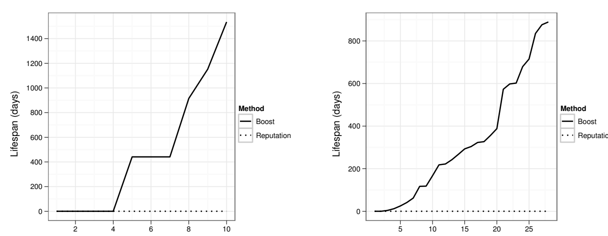
(a) Boost Rank for search term ’backpacking’

(b) Reputation Rank for search term ’backpacking’

(c) Boost Rank for search term ’photoshop’

(d) Reputation Rank for search term ’photoshop’

Figure 7: Tag search for most frequently used tag where document bookmarked 13 times or less



(a) search term ’backpacking’

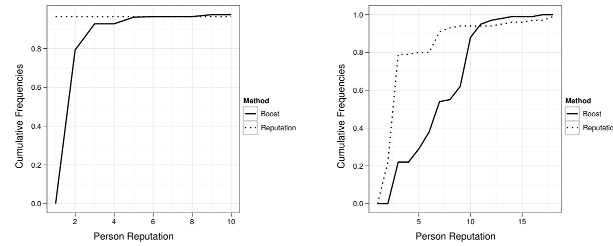
(b) search term ’photoshop’

Figure 8: Age of non-overlapping results

in Reputation Rank are high, which supports the observation that the reputation of documents with a short life-span are influenced by the users who contributed the documents. The reputation of contributing users included in the boost count distribution are also high while including a small number of users with low reputation. In figure 9 b) the maximum reputation of contributing users is slightly higher in boost rank compared to Reputation Rank. However, the overall reputation of the contributing users are higher for Reputation Rank meaning that a number of users with an above average reputation combined can have a higher influence than a small number of highly reputable users.

IV. Related Work

Social networks of the Web 2.0 content has been the subject of much recent research on mining social relationships and, most prominently, relations among tags.



(a) search term ’backpacking’

(b) search term ’photoshop’

Figure 9: Person Reputation Distribution of non-overlapping results

Heymann et. al. examined social bookmarking and tagging of data in del.icio.us [6]. The authors evaluate the utility of social bookmarking data used to augment Web search. They found that only 20% of the tags do not occur in page text or title of the pages they represent. However, they did show that social bookmarking systems provide a good reflection of changes within the underlying network, illustrating the dynamic nature of content and popularity.

Zaihrayeu et. al. attempted to calculate the trustworthiness of search results [12]. Yanbe et. al. recently developed a new page re-ranking system using social bookmark information [11]. Bao et. al. propose SocialSimilarityRank which measures the similarity between tags and SocialPageRank which accounts for the popularity among taggers in terms of a frequency ranking [1]. Hotho et. al. suggest another variation on PageRank, FolkRank, which is used to improve efficient searching via personalised and topic-specific ranking within the tag space [7]. Zanardi and Capra present a technique to re-rank results utilising the tags most associated with a given user retrieval accuracy for searches based on popular tags [13]. They also take into account the problem of searching documents based on less popular tags by expanding the user query to include tags that are found to be similar based on co-occurrence.

Golder and Huberman provides analysis of the tagging behavior and tag usage in online communities [5]. The authors provide an overview about the structure of collaborative tagging systems. Based on a small subset of the del.icio.us corpus, they investigate what motivates tagging and how tagging habits change over time. Chi and Mytkowicz investigate the dynamics of tags related to documents and shows that the information gained from a tag becomes less useful as the proliferation of use increases [2]. Based on these research results, it can be derived that time dynamics play a key role in the utility of social based ranking schemes.

V. Conclusion and Future Work

This paper has presented Reputation Ranking which measures the overall popularity of a bookmark, taking into account the timely relevance of the document. The ranking mechanism presented here is relatively simple, and involves little computational overhead by using a basic reward/decay model.

Traditional search methods apply rich-get-richer semantics, which favour old documents in the system and therefore are not able to react to upcoming trends quickly. To address this, the discriminating factor in ranking search results is not the boost count, i.e., the number of times a document has been bookmarked already, but instead a reputation metric. This paper couples the reputation of users with the reputation of the document being bookmarked in an attempt to capture the "wisdom of the crowds". As a result, reputable users, e.g., trend-setters have a greater influence on the reputation of the bookmarked documents than users with a low reputation value. Additionally, reputation is decayed over time only to be reinforced, if documents and users are constantly active. Therefore the recency of a document is an important factor when ranking for given search terms. Documents that once were popular and collected a large amount of bookmarks but have not been used recently will degrade over time. Differentiation between low-ranked documents is problematic when using traditional methods such as boost ranking. In contrast, Reputation Ranking provides a much more fine-grained ranking metric integrating a reward/decay model for both documents and users.

We believe this novel mechanism may also be applied to a single document, and the tags that are used to categorise it. As with bookmark reputation, the reputation of the tag related to the document can be evaluated. This becomes particularly useful in the case described by Chi et. al. where tag proliferation, eventually ends up degrading the meaning of the tag [2]. Heymann et. al. noted that tags were rarely applied to documents that did not contain the tag text, as a result, the mere presence of a tag is not necessarily a good representation of the extend to which the tag captures the underlying content [6]. By applying time based reputation, older tags, that may no longer be relevant, can be decayed and the presence of a tag for a corresponding document is no longer binary, but a quality measure of the tag can be harnessed.

REFERENCES

[1] Shenghua Bao, Guirong Xue, Xiaoyuan Wu, Yong Yu, Ben Fei, and Zhong Su. Optimizing web search using social annotations. In *WWW '07: Proceedings of the 16th international conference on World Wide Web*, pages 501–510, New York, NY, USA, 2007. ACM Press.

[2] Ed H. Chi and Todd Mytkowicz. Understanding the efficiency of social tagging systems using information theory. In *HT '08: Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*, pages 81–88, New York, NY, USA, 2008. ACM.

[3] Junghoo Cho and Sourashis Roy. Impact of search engines on page popularity. In *WWW '04: Proceedings of the 13th international conference on World Wide Web*, pages 20–29, New York, NY, USA, 2004. ACM Press.

[4] Scott A. Golder and Bernardo A. Huberman. The structure of collaborative tagging systems. *Technical report, Information Dynamics Lab, HP Labs*, Aug 2005.

[5] Scott A. Golder and Bernardo A. Huberman. Usage patterns of collaborative tagging systems. *J. Inf. Sci.*, 32(2):198–208, April 2006.

[6] Paul Heymann, Georgia Koutrika, and Hector Garcia-Molina. Can social bookmarking improve web search? In *WSDM '08: Proceedings of the international conference on Web search and web data mining*, pages 195–206, New York, NY, USA, 2008. ACM.

[7] A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. Information retrieval in folksonomies: Search and ranking. In *In Proceedings ESWC, 2006.*

[8] Anders Lindgren, Avri Doria, and Olov Schelén. Probabilistic routing in intermittently connected networks. In *Proceedings of the First International Workshop, SAPIR 2004, Fortaleza, Brazil, August 1-6, 2004*, volume 3126 of *Lecture Notes in Computer Science*, pages 239–254. Springer-Verlag GmbH, August 2004.

[9] David R. Millen, Jonathan Feinberg, and Bernard Kerr. Dogear: Social bookmarking in the enterprise. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 111–120, New York, NY, USA, 2006. ACM Press.

[10] James Surowiecki. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations.* New York: Doubleday, 2004.

[11] Yusuke Yanbe, Adam Jatowt, Satoshi Nakamura, and Katsumi Tanaka. Can social bookmarking enhance search in the web? In *JCDL '07: Proceedings of the 2007 conference on Digital libraries*, pages 107–116, New York, NY, USA, 2007. ACM Press.

[12] Ilya Zaihrayeu, Paulo P. da Silva, and Deborah L. Mcguinness. Iwtrust: Improving user trust in answers from the web. In *Lecture Notes in Computer Science*, pages 384–392, 2005.

[13] Valentina Zanardi and Licia Capra. Social ranking: uncovering relevant content using tag-based recommender systems. In *RecSys '08: Proceedings of the 2008 ACM conference on Recommender systems*, pages 51–58, New York, NY, USA, 2008. ACM.