

# Social Cues and Awareness for Recommendation Systems

**Punit Gupta**

Indian Institute of Technology Guwahati  
Guwahati-781039,  
Assam India.  
punitkg@yahoo.com

**Pearl Pu**

Human Computer Interaction Group  
Global Computing Center  
Swiss Institute of Technology Lausanne  
CH-1015 Ecublens, Switzerland  
Pearl.pu@epfl.ch

## INTRODUCTION

One of the major differences between navigation in the physical and the virtual world of information spaces is that the former is a social space rich in visual and other perceptual cues, be that in the form of footprints in the sand or snow, party noises we hear from an adjacent room, to dog-eared books found in libraries. Social cues and recommendations of this kind are examples for “social navigation” [3].

There are two main types of social navigation [4]. Direct social navigation takes the form of recommendations and guides, enabling people to search for information/products based on their criteria, or based on content. Indirect social navigation instead focuses on aggregating information of an entire community of users.

## MOTIVATION

We observed that information spaces like the scientific literature are still poor conduits for social navigation behavior. We might be the only person having accessed a specific research paper, but we cannot know for sure because we cannot see other people and — maybe more importantly — we are not aware of them.

Scientific literature domain is rich in information structure. There exists well established computation algorithms to calculate reputation and recommendation information, and finally digital libraries based on social navigation is gaining popularity and hence requires tools to provide the users with navigation assistance in the form of social cues as it happens in the real world navigational problems.

The goal of our system is to increase *visual* awareness of social cues existing in the information space, without making the final decisions for the user. Furthermore, the system should provide plenty of contexts for both reputation and recommendations.

## VISUALIZATION OF SOCIAL CUES

Visualization of social information enables users to easily “see” them, rather than “read” them. “Seeing” and “reading” differ in terms of the two ways the human’s low-level visual system processes information, one being automatic and other controlled processing [2][5]. However, in most existing recommendation systems, such social cues are either described in text (the NEC Research Institute ResearchIndex or Citeseer), or marked with a series of stars such as used for rating music CDs [7]. It is hard to quickly perceive reputation and recommendation information in text. Or, a simple visualization mechanism does not scale to complex ranking.

## REPUTATION AND RECOMMENDATION

Reputation means how the rest of the world feels about a particular item. For example, GOOGLE calculates the reputation of a page from the number of links pointing to it and displays the most *popular* item on the top of a linear list [1].

Recommendation, on the other hand, is computed information, helping users to find similar items and related items based on what information the user has already obtained, or his profile. This is a useful way to guide a newcomer. He can follow the guidelines left by previous users. Imagine yourself in a shop, and you want to buy a CD player. Since you have never bought it before, you do not have any information about it. So recommendation of your friend, shopkeeper or any customer may be an immense help to you. Visualization work in this area has been very limited, mostly employing textual labeling such as “hot” or “popular item”.

Researchers from a number of areas have been working on the calculation and use of reputation and recommendation. For a survey, please refer to [9][6]. We focus our investigation on visualizing reputation and recommendation in scientific literature, sometimes called the scientometrics.

## REPUTATION IN SCIENTIFIC LITERATURE

What counts in reputation in a domain must first be analyzed in the user task context. We conducted a survey on the factors influencing a literature searcher among a group of 12 graduate students in their initial phase of Ph.D. studies at the Swiss Federal Institute of Technology (EPFL). They have diverse areas of specialization such as human computer interaction, databases, audio-visual communications, mobile communications, mathematics and physics. They are from very different demographic backgrounds such as northern America, China, India, Mexico, Romania and Yugoslavia. This group has been selected because they frequently perform literature search, and yet do not have enough knowledgeable in a domain. They rely very much on reputation and recommendation information provided by the others.

Subjects were given a list of factors that we believe to influence the decision-making process in scientific literature search, and they were asked to rate the factors according to their preferences. We studied the feedback received from the users and sorted the users' preferences in terms of how they view a paper as worthy of pursuing. They are:

1. In which journal was it published,
2. Popular authors in the field,
3. Comments of other users about the paper,
4. Rating of citing documents,
5. Number of citations,
6. Rating of the paper,
7. In which conference was it presented,
8. Reputation of the research lab involved,
9. Year of publication,
10. Number of people who read it.

As it turned out, there are many factors comprising reputation. As many of them are not available in digital libraries, we focus on modeling the following factors:

- Popular authors in a field,
- Number of citations of an article
- Year of publication.

However, displaying a 3D of information can be challenging for web browsers and confusing for users. Furthermore, a multi-bar chart display takes much space and cannot be used to compactly display many items for comparison. For the particular digital library (Citeseer) for which we are designing the interface, the popularity of an author (how many times he is cited) is displayed separately from the articles that he/she has published. We have therefore decided to employ a 2-dimensional representation to visualize the number of citations and the year of publication.

## Visual mapping of the scales of reputation

A close observation of patterns and ranges of citations suggests that in most cases there is an unusual high density of papers in the low ranges of citation and the wide gap between low and high number of citations [7]. Therefore, if we represent the citation of papers on a linear scale using for example font size or a series of stars will cause visual anomalies: there will be a large amount of perceptual space wasted in the middle range since few papers' citation numbers fall in that range; some papers will take too much space to visualize while others will be almost invisible. Furthermore, papers which were published recently tend to have few citations, and will be unjustly de-emphasized. Thus it calls for some non-linear mapping techniques between the actual citation value and what can be visually displayed. The mapping will result in a range of 10 distinct visual ranges. We have opted for the slow-in slow-out method [8], which magnifies the highly dense areas, and de-magnifies in the low distribution area.

To map these scales to visual values, a two-dimensional display is employed where the documents are ordered by their year of publication, and each document is visually rendered to show its citation number. After several explorations, we have finalized upon five designs to map the scales to the visual renderings. 1- Font sizes, 2 - Stars, 3- Stars + details, 4- Font sizes + shading (weight), 5- Font sizes + shading (weight) + details.

We conducted an empirical study to test users' acceptance of these five designs and determined the most effective non-linear mapping of reputation scales. Results show that the design employing 'Font sizes + shading (weight) + details' is the most effective method to represent reputation of research papers (Fig 1). For details on the user study, please refer to [7].

<b>Oxford University Press</b>	<b>Dawkins</b>	<b>1976</b>
<b>Information retrieval</b>	<b>vanRijsbergen</b>	<b>1979</b>
<a href="#">Journal of Experimental Psychology: Learning (context)</a>	Anderson, Pirolli	1984
<a href="#">Princeton University Press (context)</a>	Stephens, Krebs	1986
<a href="#">Phase transitions in artificial intelligence systems (context)</a>	Huberman, Hogg	1987
<a href="#">Human memory: An adaptive perspective (context)</a>	Anderson, Milson	1989
<a href="#">The adaptive character of thought (context)</a>	Anderson	1990
<a href="#">Identifying appropriate hypertext structures (context)</a>	Botafogo, Schwabman	1991
<a href="#">An Object-Oriented Architecture for Text Retrieval</a>	Cutting, Pedersen et al.	1991
<a href="#">Proximate ecology and human behavior (context)</a>	Smith	1992
<a href="#">Information visualization using 3D interactive animation (context)</a>	Robertson, Card et al	1993
<b>Rules of the mind (context)</b>	<b>Anderson</b>	<b>1993</b>
<a href="#">Indexing the Internet (context)</a>	Teuber	1995
<a href="#">Results from the Third WWW Survey (context)</a>	Pitkow, Johnson	1995
<a href="#">Information foraging in information access environments (context)</a>	Pirolli, Card	1995
<a href="#">Visualizing complex hypermedia networks through multiple hierarchies</a>	Mukherjee, Foley et al	1995
<b>Characterizing browsing strategies in the World-Wide Web</b>	<b>Catledge, Pitkow</b>	<b>1995</b>
<a href="#">Father browsing communicates the topic structure of a very large WWW site (context)</a>	Pirolli, Schank et al.	1996
<a href="#">The WebBook and the Web Forager: An Information Workspace</a>	Card, Robertson et al	1996

Figure 1: Design using 'font size + shading + weight'

## RECOMMENDATION

Our case of recommendation deals with providing users with social cues to examine other documents of interest based on the current document. At Citeseer, five categories

ries of recommended documents are displayed: *cited-by* documents (those documents citing the current article), *active bibliography* (related documents at the sentence level), *similar documents based on text*, *Related documents from co-citation*, and *citations* (documents cited by the current article).

### Metaphor design for recommended items

We explore the different metaphor designs to facilitate easy recognition of the types of recommendation, and at the same time easy *recalling* of the icons to facilitate task performance after a period of interruption (such as after a vacation or a period of inactivity). We designed two sets of icons for each type of recommendations used in Cite-seer and conducted a study with the users to select the most appropriate icon set for our case [7]. Table 1 shows the two sets of icons for each type of recommendations used in Cite-seer. Our informal user studies showed that icons from Icon 1 are slightly preferred by our users.

**Table 1: Two sets of icons for different categories of papers**

Category	Icon 1	Icon 2
Cited-by documents		
Cited document (active bibliography)		
Similar documents at sentence level		
Related document from co-citation		
Similar document based on text		

We also explored two different visual layouts. In the first layout (L1), each recommended paper is marked with an icon, while in the second layout (L2), each category of papers is marked by a single icon, thus reducing the number of icons appearing on the screen. Our user study showed that users prefer L2 to L1 and to the original Cite-seer layouts. For details on the user study please refer to [7].

### CONCLUSION

In this paper, we have argued for a visualization approach to enable users perception of social cues and augment their awareness of each other in social navigation environments. For a particular case study where scaling issue is most important, we have explored various visual emphasis techniques to effectively represent reputation and recommendation in scientific literature search. Due to space limitation, please refer to [7] for the redesign of the interface for Cite-seer.

### REFERENCE

- [1] Brin S. and Page L. (1998). *The anatomy of a large-scale hypertextual Web search engine*. In Proceedings of the Seventh International World Wide Web Conference, Brisbane, Australia. Elsevier.
- [2] Card S.K. (1996). *Visualizing Retrieved Information: A Survey*. IEEE Computer Graphics and Applications, 63-67.
- [3] Dieberger, A., Supporting Social Navigation on the World Wide Web. International Journal of Human-Computer Studies, 1997. 46: p. 805-825.
- [4] Dieberger A. Where did all the people go? A collaborative Web space with social navigation information.
- [5] Janecek P. and Pu P. (2002). *A Framework for Designing Fisheye Views to Support Multiple Semantic Contexts*. International Conference on Advanced Visual Interfaces (AVI '02), ACM Press, 2002.
- [6] Marsh S. (1994). *Formalizing Trust As a Computational Concept*. PhD thesis, University of Stirling, U.K.
- [7] Pu P. and Gupta P., Visualizing Reputation and Recommendation in Scientific Literature. Submitted in HCI International 2003.
- [8] Sarkar M. and Brown M.H. (1992). *Graphical Fish-eye Views of Graphs*. In Proceedings of ACM Conference on Computer Human Interaction.
- [9] Yu B. and Singh M.P. (2000). *A social mechanism of reputation management in electronic communities*. In Proceedings of Fourth International Workshop on Cooperative Information Agents.