

A Refinery of an Internet-Based Search Tool – Exploring Perceptions from Information Systems Practitioners

Michael Chau¹, Ivy Chan²

¹School of Business, Faculty of Business and Economics
The University of Hong Kong, Pokfulam, Hong Kong

E-mail: mchau@business.hku.hk

²Department of Decision Science and Managerial Economics
The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong

E-mail: ivychan@baf.msmail.cuhk.edu.hk

Abstract: In today's dynamic business environment, the capability to understand the needs and responses of stakeholders is critical as such management can devise effective plans and course of actions for long-term strategies. The advancement of the internet and its related search tools such as Google, has helped management to collect various information in order to keep abreast of time of business environment and their business communities. Yet, the management has encountered problem and uncertainty on the information quality such as content, currency, accuracy and presentation. Therefore, a new backlink search tool, namely Redips was developed. This tool has been tested by students subjects with satisfaction. However, in order to gain acceptance from the management (i.e., the end-users), the perceptions and expectations on the tool should not be overlooked. This study presents a wide range of perceptions from a group of Information Systems practitioners on the potential use of Redips.

I. Introduction

The World Wide Web presents significant opportunities for business intelligence analysis as it can provide information about a company's external environment and its stakeholders. Traditional business intelligence analysis on the Web has focused on the simple keyword searching. Recently, it has also been suggested that the incoming links, or backlinks, of a company's Web site (i.e., other Web pages which have a hyperlink pointing to the company of interest) can provide important insights about the company's "online community". Analysis of these communities can provide useful signals for a company or information about its stakeholder groups, but the manual analysis process can be very time-consuming for business analysts and consultants. In this paper, we present a tool called Redips that integrates automatically backlink meta-searching and text mining techniques to facilitate users in performing such business intelligence on the Web. We also report a focus group study involving IS practitioners to study the potential uses and user perception of the proposed tool. The rest of the paper is structured as follows. Section 2 reviews the importance of Web communities in business intelligence analysis. Section

3 presents the proposed tool and its architecture. In Section 4 we discuss our methodology and the focus group study. The preliminary results are presented in Section 5. Finally we conclude our study and suggest future research directions in Section 6.

II. Web Communities

The Internet has many well-known explicitly defined communities – groups of individuals who share a common interest, together with the Web pages most popular amongst them [10]. The Web communities consist of the following stakeholders of the firm: customers, suppliers, competitors, regulators, employees, educational institutions, court and legal institutions, financial institutions, stockholders, public-interest groups, labor unions, political parties, federal, state, local governments, etc. [11]. The stakeholders listed here can be classified into two categories: explicit and implicit Web communities.

Explicit communities are the communities that can be easily identified on the Internet. Kumar et al. [8] discussed the Porsche newsgroup as an example of explicit community of Web users interested in Porsche Boxster cars. Such communities are often found in resource collections in Web directories such as the Yahoo directory. Explicit communities are easy to be identified and analysts can simply use manual method to find a firm's explicit communities by browsing the firm's newsgroup or the category in which the firm belongs to in directories like Yahoo on the Internet.

Implicit communities are relatively more difficult to be found using a manual browsing method. According to Kumar et al. [8] implicit communities refer to the distributed, ad-hoc and random content-creation related to some common interests on the Internet. These pages often have links to each other, but the common interests of implicit communities are sometimes too narrow and detailed for the resource pages or the directories to develop explicit listings for them. As a result, it is more difficult to find the implicit communities of a firm. In identifying the explicit and implicit communities of a firm, it is reasonable to assume that the content pages created by these communities would provide hypertext links back to the firm's homepage for reference [10]. Therefore, in order to find a firm's online communities, it is necessary to find the Web pages that have

hyperlinks pointing to the firm's URL, i.e. the inbound links of the firm's Web site.

The identification of Web communities, irrespective of explicit communities or implicit communities, is important to the strategic planning process. The strategic planning process consists of five steps [3], namely, mission and objectives, environment scanning, strategy formulation, strategy implementation, evaluation and control. The extraction of Web communities is classified to the environmental scanning step in the strategic planning process. The information would be used for the analysis of the firm's industry for evaluating entry barriers, suppliers, customers, substitute products and industry rivalry.

III. Redips

Our proposed system is called Redips. The term "Redips" is the reverse spelling of the term "Spider". Traditional Web spiders (crawlers) search the Web by following outgoing links. However, in our proposed tool, we search based on the "incoming links" of a page. These incoming links can help identify the Web communities of the firm of interest.

Redips has been implemented based on the MetaSpider system developed in our previous research [6]. The main modules include the user interface, spider, Arizona noun phraser, and self-organization maps. User interface is the first point of contact between the user and the system. Spider fetched the URLs returned from those search engines. Arizona noun phraser is a natural language processing tool to do the key phrase extraction from Internet text. Self-organization maps visualize the concepts in a two-dimensional map. The technical details of the modules are discussed in another paper [5]. Here, we will present a search session with the tool from the user's perspective.

When using Redips, a user should first enter the Web site to be analyzed and the backlink search engines to be included. A sample user session with Redips is shown in Figure 1. In this example, the Web site entered is <http://www.ibm.com/>, the homepage of the IBM Web site. Optionally, the user can enter the keyword(s) to be included in the returned Web pages. The user may also specify some other search options. In this step, the user can define the intelligent analysis objectives, e.g. the firm, information source, topic, in the analysis process. After starting the search, multiple threads will be generated to start getting Web pages from the Internet. The URLs returned by the search engines will then be displayed. The user can browse these pages for exploratory, preliminary research in this step.

IV. Research Methodology

As discussed above, there are a number of new attributes embedded in Redips, which facilitate information search and add values to the management in strategic planning. Yet, the attributes of Redips perceived and developed by the system developers may be different from those of the business managers. In order to provide an effective and practical tool,

the views and comments from IS practitioners who are rich in IS knowledge and currently engage in business planning, leadership and strategic analysis are examined. The current study adopts focus group study (as an effective research method) to dig in and explore various perceptions from the IS practitioners, as such to enrich our understanding of the practicability and usefulness of Redips.

Focus group discussion is a qualitative method to trigger and solicit extensive pool of opinions that capture the respective dimensions of the domains or topics to be addressed or investigated, particularly newly emerged and illuminative concepts [7,9]. Generally speaking, focus group is a method conducive to communication through a two-hour back-and-forth discussion and interaction among a small group of individuals (usually six to ten persons). A moderator is enlisted to facilitate the discussion (e.g., to ensure topics and issues are conversed, commented upon and incorporated). In our current study, one focus group discussion was conducted. The group was composed of six practitioners (in their mid-thirties) with an average of 12 years of industry experience. They discussed on the search approaches, pros and cons of commonly used internet-based search tools. Thereafter, they engaged in a 30-minute hands-on of the Redips and explicated their opinions from various views. Large amount of broad-based ideas in participants' own words, while relatively inaccessible using other research methodologies can be obtained [4]. The coding and analysis of the data, thereby can be used to develop thematic scheme to reveal the scope, details and multiple facets of the constructs or ideas in proposed work [1, 12].

The focus group study serves as the beginning stage of research. We intend to use a mix of techniques to make sense of the qualitative data collected from the practitioners. A research plan with three stages is detailed as follows:

Stage One. In order to understand how users perceive and evaluate new systems, a reference to existing IS literature was undertaken. The pertinent literature is used as a foundation to outline the topic agenda used in the focus group sessions, while refrained from the actual discussion.

Stage Two. A series of focus group sessions is planned to be conducted. The first focus group discussion was held in August 2005 which provided a lot of useful feedbacks of Redips. More focus group sessions are planned, involving business executives from non-IS functional areas (such as sales and marketing sections in their organizations) of which the executives are expected to use the Internet-based tool frequently in their daily tasks.

Stage Three. Data collected (in handwritten notes and audio tapes) from different focus group sessions will be firstly compiled with the assistance of the computer. The results will thereby be compared and assimilated. A validated and accuracy check was conducted to compare the summary and the original transcripts. The adoption of thematic categorization [1,2,12] involves the articulation and construction of potential themes and facets of constructs that correspond with the research questions. The potential themes

were considered as a reflection of the attributes or dimensions captured under each construct. To ensure no interpretive perspectives would be lost, the transcripts (recorded in word documents) were sorted and allocated to the emerged themes using a cut and paste technique. Throughout iterative rounds of discussion and reconciliation, the discrepancies of interpretation were resolved to devise a theme summary. The theme summary includes coding systems, themes titles, descriptions, relevant examples and indicative keywords. To ascertain the effectiveness and accuracy of the system evaluation theme summary, further comments were sought from other experts (who are not involved in any stage of the current study).

V. Preliminary Results

The first focus group session was held, with large amount of ideas and opinions were sought. In essence, the practitioners are satisfied with the existing tools, in regard to its low cost (usually free of charge) and ease of use. However, the information quality in terms of amount, currency and presentation has presented limitations to their business decisions. The participants reveal that Redips can add value to their managerial tasks, as it streamlines the information search, and enables effective assessment of the Web

communities and stakeholders. A common assertion was sought that they will use Redips in addition to the existing search practice.

Some key preliminary findings are summarized as follows:

The practitioners had named Yahoo and Google as two commonly used Internet-based tools in their daily work. There are some extraordinary occasions that they may seek information from other search tools (such as www.altavista.com) given the two common tools did not provide adequate information for their purposes.

To evaluate an effective and desirable system, nine attributes are assimilated from the opinions sought in the focus group discussion (Table 1). Redips' performance shows a good match and desirable attributes with those expected by the IS practitioners. The attributes are ranked according with the number of respondents who solicit the respective aspect. Therefore, a higher the number in the bracket indicates a more common issue being conceived in the discussion session. It should be noted that one of the advantages of focus group discussion is to provide analytical generalization of what participants think and perceive, therefore, it is not meant the higher the number of a respective attribute as more important one.

Table 1. Desirable Attributes of an Internet-based Search Tool and Redips

Desirable Attributes	Details and Descriptions	Redips
Time (6)	Short and prompt search time	Quick response and fetching time
Cost (6)	Free of charge	Free of charge
Search results (6)	Links should be ranked with relevance and currency of the information	Web pages are fetched in real-time to avoid outdated indexes
Learning efforts (5)	Minimum, ease of use without attending formal training	Easy to use, logical and order search process
Technical resources (4)	No additional software be installed	Requires installation
Interface (4)	User-friendly and attractive	Practical and clean, user-friendly buttons
Language (3)	Multiple language	Limited to English
Personalization (2)	Customization of specific search requirements, prior search records	Three search engines (including Google, MSN and AltaVista) can be previewed at the same time; phrase selection to remove unrelated pages
Assistance (1)	Help, search tips and data recovery	Help menu available

Contrast to our result-oriented system design, the IS practitioners who are experienced in system development and engage in business planning share similar perception and valuation criteria that are close to those of end-users'. Most of the participants showed empathy on the business managers' pressure in time management. A system analysis manager stated that the tool should be time saving, with results displayed with highly relevant, clear, easily be

comprehended and useful content. Two other participants supported his views as they claimed the business development managers in their organizations are "extremely" hard working as they usually work for more than 54 hours per week, thus they cannot afford information search process that "wastes" their time. Therefore, they stated that the search process should be time-effective, with minimum search and loading time. They commented further

that long loading time will increase the tendency to abort the search process, thus, unfavorably affect managers' time management.

While discussing the approaches in search of information or business intelligence, the practitioners stated that their search approaches vary from task to task. One IS department assistant manager stated that if he looks for new business opportunities that he has little experience about it or the request is made on ad hoc basis, he may seek information exhaustively until the relevance of the search results diminishes. He explained "I will seek help from experienced staff within organization. Yet, sometimes good business ideas are not generated from internal people as they are indulged to the existing success. Those competitors, discussion groups provide lots of comments to our existing services, despite their words are unpleasant. Listening to our customers can let us know our inefficiency and what we need to do more." Otherwise, if he handles regular review of market positioning, he will approach a satisfactory manner. He stated "my personal experience is rich and validated in my past success. Therefore, I know when and where to get the relevant information for routine tasks, and know when to stop."

In regard to the propagation of Redips, a senior system engineer pointed out that a good design does not make the system successful and well accepted by users. He stated, "Our colleagues have a strong social network which shared lots of latest information and users' feedback on various tools." Therefore, he suggested "the success of Redips is not a mere concern of physical design and functionality; rather it is determined by the consumers who foster a good perception at the beginning. If the comment is positive, the word-of-mouth will leverage users' interest and attention to Redips." A senior software development manager put forth the opinion and stated that management should not underestimate the power of management commitment and network effect. She gave an example that a new software system was well accepted in her organization, despite the brand equity of that software system is not high. It is because the management has shown a strong support, clear vision and commitment to deploy across all departments. Middle management also plays an important role to monitor the roll out of the system. Furthermore, she recalled "most of our colleagues believe that they should get on the same board (i.e. adopt the new system) in order to become one of the majority."

VI. Conclusion

The internet-based backlink search tool, Redips provides an alternative way to assess various web communities and infer business intelligence to managers. In order to enhance its practicability and usefulness to business managers, the current study adopted focus group method to drill in various

perceptions and views from a group of IS practitioners. In essence, they find Redips and assert its value in business planning and decision making. Our present study will be continued and extended as proposed the three-stage research tasks. The nine attributes and specific findings about managerial problem solving approaches will be also incorporated as such Redips can be widely adopted.

Acknowledgement

This research has been supported in part by the following grants:

- HKU Seed Funding for Basic Research, "Searching the World Wide Web Backwards for Business Intelligence Analysis," January 2005 – December 2006.
- HKU Seed Funding for Basic Research, "Using Content and Link Analysis in Developing Domain-specific Web Search Engines: A Machine Learning Approach," February 2004 – January 2006.

We would like to thank the IS practitioners for their participation in our study.

References

- [1] Blackburn, R., and Stokes, D. "Breaking Down the Barriers: Using Focus Groups to Research Small and Medium-Sized Enterprises," *International Small Business Journal* (19:1), 2000, pp. 44-67.
- [2] Boyatzis, R. E. *Transforming Qualitative Information: Thematic Analysis and Code Development*, Sage, Thousand Oaks, CA, 1998.
- [3] Bradford, R. W., Duncan, P. J., and Tarcy, B. *Simplified Strategic Planning: A No-Nonsense Guide for Busy People Who Want Results Fast!* Chandler House Press, 1999.
- [4] Caldeira, M. M., and Ward, J. M. "Understanding the Successful Adoption and Use of IS/IT in SMEs: an Explanation from Portuguese Manufacturing Industries," *Information System Journal* (12:2), 2002, pp. 121-152.
- [5] Chau, M., Shiu, B., Chan, I., and Chen H. "Automated Identification of Web Communities for Business Intelligence Analysis," *The Fourth Workshop on E-Business (WEB)*, Las Vegas, Nevada, USA, December 2005.
- [6] Chen, H., Fan, H., Chau, M., and Zeng, D. "MetaSpider: Meta-searching and Categorization on the Web," *Journal of American Society for Information Science & Technology* (52:13), 2001, pp. 1134-1147.
- [7] Krueger, R. A. *Developing Questions for Focus Groups: Focus Group Kit 3*, Sage, 1998.
- [8] Kumar, R., Raghavan, P., Rajagopalan, S., and Tomkins, A. "Trawling the Web for Emerging Cyber-communities," in *Proceedings of the Eight International World Wide Web Conference*, 1998.
- [9] Morgan, D. L. *Focus Groups as Qualitative Research: Qualitative Research Methods Series 16*, Sage, 1988.
- [10] Reid, E. O. F. "Identifying a Company's Non-Customer Online Communities: a Proto-typology," in *Proceedings of the Hawaii International Conference on System Sciences*, January 6-9, Big Island, Hawaii, 2003.
- [11] Schermerhorn, J. R., 2001. *Management*. John Wiley & Sons, Inc.
- [12] Stewart, D. W., and Shamdasani, P. N. *Focus Groups: Theory and Practice*, California: Sage, 2000.

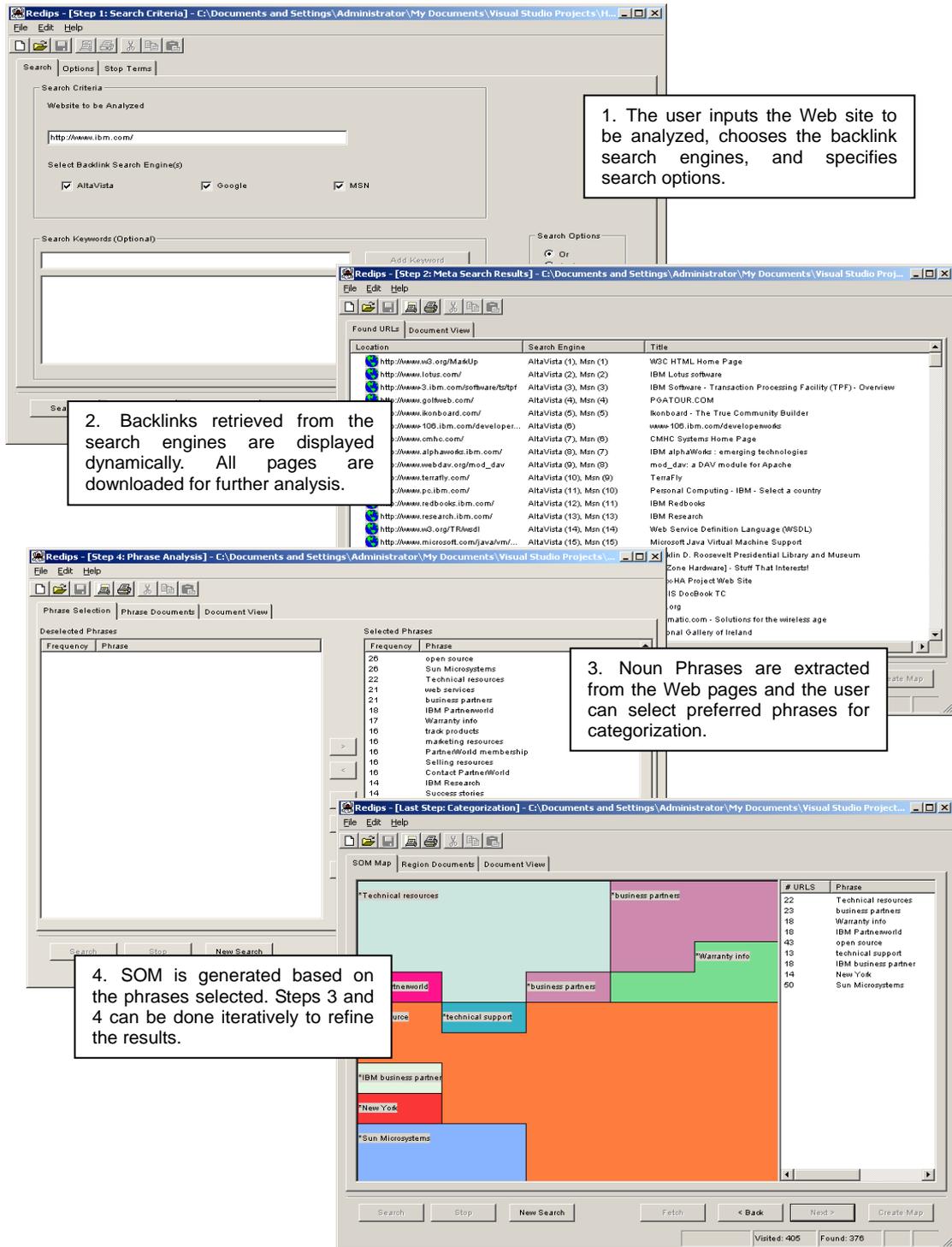


Figure 1: Example of a User Session with Redips