Encyclopedia of Multimedia Technology and Networking

Second Edition

Margherita Pagani
Bocconi University, Italy

Volume III
O–Z

Information Science
REFERENCE

INFORMATION SCIENCE REFERENCE

Hershey • New York

Director of Editorial Content: Kristin Klinger
Senior Managing Editor: Jennifer Neidig
Managing Editor: Jamie Snavelly
Assistant Managing Editor: Carole Coulson
Cover Design: Lisa Tosheff
Printed at: Yurchak Printing Inc.

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue, Suite 200
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: <http://www.igi-global.com/reference>

and in the United Kingdom by
Information Science Reference (an imprint of IGI Global)
3 Henrietta Street
Covent Garden
London WC2E 8LU
Tel: 44 20 7240 0856
Fax: 44 20 7379 0609
Web site: <http://www.eurospanbookstore.com>

Copyright © 2009 by IGI Global. All rights reserved. No part of this publication may be reproduced, stored or distributed in any form or by any means, electronic or mechanical, including photocopying, without written permission from the publisher.

Product or company names used in this set are for identification purposes only. Inclusion of the names of the products or companies does not indicate a claim of ownership by IGI Global of the trademark or registered trademark.

Library of Congress Cataloging-in-Publication Data

Encyclopedia of multimedia technology and networking / Margherita Pagani, editor. -- 2nd ed.
p. cm.

Includes bibliographical references and index.

Summary: "This publication offers a research compendium of human knowledge related to the emerging multimedia digital metamarket"--Provided by publisher.

ISBN 978-1-60566-014-1 (hardcover) -- ISBN 978-1-60566-015-8 (ebook)

1. Multimedia communications--Encyclopedias. I. Pagani, Margherita, 1971-

TK5105.15.E46 2009

621.38203--dc22

2008030766

British Cataloguing in Publication Data

A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this encyclopedia set is new, previously-unpublished material. The views expressed in this encyclopedia set are those of the authors, but not necessarily of the publisher.

If a library purchased a print copy of this publication, please go to <http://www.igi-global.com/agreement> for information on activating the library's complimentary electronic access to this publication.

Scanning Multimedia Business Environments

S

Sören W. Scholz

Bielefeld University, Germany

Ralf Wagner

University of Kassel, Germany

INTRODUCTION

The term *environmental scanning* (ES) refers to the way in which managers study their relevant business environment. More precisely, we define *ES* as both looking for and looking at information available in the business environment. These activities embrace all domains of gathering facts from external sources like *competitive intelligence* (CI) and market research, but take a more holistic, integrative perspective by considering consumers, competitors, and the technological developments in same industry and different industries as well (Scholz & Wagner, 2006). Choo, Detlor, and Turnbull (2001) give an overview of ES and empirical findings supporting the importance of ES activities in business organizations. ES helps managers foresee favorable as well as unfavorable influences and initiate strategies that enable their organizations to adapt to their business environments.

This article:

- outlines three perspectives on the challenges for environmental scanning arising from information dissemination by multimedia,
- discusses opportunities for ES on the World Wide Web (WWW), and
- exemplarily describes two software solutions for ES in multimedia.

The remainder of the article is organized as follows: the next section discusses challenges for ES arising from the increase of multimedia technologies in the business environment. Then, we discuss the impact of the WWW on ES. Subsequently, we present the state of the art in ES practice as well as the supporting software. The article concludes with a brief outline of future challenges in the field of ES.

ES IN MULTIMEDIA: CRAFTSMANSHIP, ART, OR SCIENCE?

Craftsmanship

Referring to Jauch and Glueck (1988), the external environment consists of the following six areas: (1) customers, (2) suppliers, (3) competition, (4) socioeconomic, (5) technological, and (6) governmental. As depicted in Figure 1, these sectors make up the firmament of an organization's information environment.

In the figure, the sectors are ordered with respect to their affinity to the industry environment on the left-hand side and the global *business environment* on the right-hand side. On the bottom plate, the internal information sectors are illustrated. The relevant internal business environment comprises research and development, market research, basic engineering, cost controls, financial management, and controlling departments (Garg, Walters, & Priem, 2003). Major aims are cost control and operational efficiency, but information fragments may be found in all internal media, particularly in the intranet, in-house presentations and protocols of all kinds of meetings. All tasks appear to be well ordered and, thus, sophisticated *ES* activities should be a good piece of artisanship. However, the artisan has to cope with two problems: the wide array of relevant topics and the diversity of information sources. Consequently, there is constant competition for the manager's limited attention among different topics, information sources, and fragments. Accordingly, scanning is a challenging task because a broad range of internal and external sources have to be exploited, data in different (often ill-specified) formats have to be combined, and the topics as well as the information sources of interest cannot be exhaustively described, *a priori*, but rather emerge during the scanning activities.

Art

ES aims to improve short-term and long-term planning and should lead to a better understanding of external changes. However, in the real world of management, *ES* is still far from being recognized as a structured task applied in daily managerial practice, for various reasons. First, this might be attributed to the perception of managers that systematic *ES* is user unfriendly (because of its quantitative methods), too complex (and thus simultaneously oversimplified, because of reducing to scope to very few of the relevant variables) and, therefore, might make them lose ground (Day, 2002). This problem becomes even more serious in the context of *ES* activities relying on multimedia, because the information is typically not presented in a well-structured and easily accessible manner. Professional tools for managing the *ES* process and systematizing these information sources are still not widespread in practice (Benczúr, 2005). Standard search engines, such as Google or Ask Jeeves, seem to be a great help in retrieving meaningful, relevant information at first glance, but do not organize the search results in a serviceable manner. Like art, *ES* requires connotations different from simple Web searches.

Science

The scientific perspective on *ES* in the *business environment* is dominated by the aim of identifying *weak signals* heralding significant changes in the industry structure or the relevant technology (cf. Martino, 2003, and Scholz & Wagner, 2006, for systematic literature reviews). In the early stages, the signals are small, subtle hints that are hardly distinguishable from the background noise. However, the earlier the organization detects these *weak signals*, the more time it has to successfully align the strategic decisions to emerging, forthcoming changes in the business environment. The central task of *ES* is to provide sound methodologies for the successful detection of relevant changes heralded in the information environment.

Modern computer technologies can help to identify the maturity of signals by means of data mining and bibliometric analysis of textual information sources. Intranet technologies provide promising solutions for managing knowledge repositories, and allow quick access to the sparse and distributed internal knowledge on the *external environment*. The most relevant

characteristics of the information fragments provided by multimedia are their vagueness and the absence of links to other information that would make a consistent and meaningful picture. Thus, linking information fragments from different sources is the main challenge for the identification of *weak signals*.

The problem becomes more difficult if the prior knowledge of the decision maker is included in the problem formulation. Clearly, individuals have different knowledge of topic domains. They might be experts in very few domains, but these are not necessarily the relevant domains for grasping information fragments. The *information foraging theory* (Scholz & Wagner, 2006) provides us with a formal foundation for both modeling information seeking behavior and implementing innovative tools for gathering and combining information spreading over many media objects. An important element of this formal framework is the information infrastructure that captures the prior knowledge already gained with respect to an interesting topic. All incoming information fragments are assessed with respect to this prior knowledge, and the most interesting ones are selected for further consideration. In an empirical evaluation this methodology has been found to work well, even if errors are included in the information infrastructure (Scholz & Wagner, 2005).

WHAT A SIGHT: SCANNING THE WWW

The Internet—and in particular the WWW—provides the manager with excellent external information. Nowadays, the WWW consists of several billion documents that are largely designated for cost-free use. Trends in business, science, society, and politics are digitally announced on the WWW long before their consequences are observed in the real world (Decker, Wagner, & Scholz, 2005). McGoangle and Vella (1998) argue that 90% of all information needed by a company to make critical decisions is either already public or can be systematically developed from public data. As the WWW is both a publishing medium and an indispensable element of daily communication, almost all up-and-coming real-world phenomena are discussed in the virtual reality of the WWW. Therefore, Choo et al. (2001, p. 161) refer to the Internet as a “social information space.”

Figure 1. Arenas of an organization's environmental scanning activities

LOSING ONE'S HEAD: SCANNING THE WWW

However, due to the massive information supply—Google already indexes more than eight billion Web pages—efficiently searching this information space requires a considerable understanding of ways to manage the selection of information fragments primarily because the WWW opens up a mind-blowing volume of ill-structured information. Search engines provide fundamental support in detecting relevant information, but are far from being adequate tools for *ES*. The dynamic and largely unregulated nature of the WWW is making it increasingly difficult to locate relevant information at reasonable costs (Chen, Chau, & Zeng, 2002).

Search engines have several drawbacks in presenting retrieved information. A search engine only proposes simple links that have to be followed manually. Furthermore, these search engines frequently provide too many irrelevant responses. Fine-tuning or the use of advanced search engines takes time and requires special knowledge. One aim of technical advances in *ES* systems is to develop sensitive and context-dependent reductions of the massive information overload that makes up the environment in which modern businesses have to operate.

From a practitioner's perspective, systematizing the *ES* process in an integrated software is clearly one of the most pressing future challenges. Both increasing efficiency and clarifying the information needs, and thus triggering the *ES* process in a dynamic, volatile,

and unstructured environment such as the WWW, is needed.

THE YIN AND YANG OF ES

As described by the allegory of yin and yang, the crucial problem in *ES* processes is to knit the abilities of the human mind and the complexity of modern business environments. Multimedia processing systems are clearly the most promising tools for supporting the environmental scanning process. Managers' mental models not only facilitate the link between managerial decision-making and *ES*, but also define the criteria *ES* systems have to meet (Scholz & Wagner, 2006). A sound understanding of the user's way of thinking and his or her underlying information needs allows a customization that will increase the usability of the system. Subsequently we discuss software prototypes designed to cope the differences between the users' imagination of the business environment und the wide spread of information fragments.

TWO PROTOTYPES: MASTERSCAN & CI SPIDER

The MasterScan, developed by Liu (1998), is a prototypic *software agent* that exemplarily executes scanning tasks for the pulp and paper industry. This system visits previously defined Web pages and gathers relevant

Figure 2. Screenshots taken from the CI Spider (with kind permission of Hsinchun Chen)

information from these sources. The MasterScan is just able to extract information from pre-selected sources by comparing the content of these Web sites with the results from previous visits. Therefore, it is primarily a monitoring tool that is not designed to browse for new sources and detect information that is weak in the sense that basic ideas and concepts about these data have not yet been elaborated.

In a more recent approach, Chen et al. (2002) have developed a tool for *CI* on the Web that is also based on *software agent* technology. The user has to specify a starting URL and provide the keywords in order to evaluate the links and the documents found by the Spider. The CI Spider tries to avoid the principal shortcomings of classic search engines by generating its own structure for the retrieved documents and displaying them in a two-dimensional chart “on the fly” by means of a *self-organizing map* (SOM) approach. The CI Spider outranges standard Web search engines—for example, by the following graphical representation of results as well as by sophisticated extraction and visualization of noun-phrases included in retrieved documents by means of part-of-speech tagging and linguistic rules. Accordingly, empirical results show that the CI Spider provides significant better performance results with respect to precision, recall, and ease-of-use than a standard search engine (Chen et al., 2002). However, it is still based on a standard keyword search. The only possibility to realize completely undirected search activities is given by providing just a starting URL for a Web site of interest. In Figure 2, this has been conducted for the ehotelier newsletter (see www.ehotelier.com).

As shown in Figure 2, the user interface in the left screenshot is clearly structured, and only very few inputs have to be provided for a search. The right screenshot depicts the CI-Spider illustration of the SOM grid. This visualization can be easily modified by including different noun-phrases in the search results. However, the user still needs a rather complete picture of the topics and phenomena he or she wants to scan or monitor on the Web. The detection of completely new phenomena in the business environment is not featured. Due to the limitations of the classic SOM grid, the results for such a broad search are limited. Nevertheless, the CI Spider yields reasonable results if precise key words are provided by the user.

EXPLICATING THE INFORMATION STRUCTURE

An innovative approach to *ES* on the WWW has been introduced by Ontrup, Ritter, Scholz, and Wagner (submitted). This approach specifically takes into account different scanning activities that arise from the current mental state of the manager. The system supports the *ES* process in various ways: The gathered information is first used to train a hierarchically growing hyperbolic map (H^2SOM), which is an extension of the classic *SOM*, in a completely unsupervised way. The resulting *information structure* can be interactively explored by mapping the nodes of the H^2SOM on a Poincaré Disk (see Ontrup & Ritter, 2006, for details) in a similar fashion to the hyperbolic tree

Figure 3. Two screenshots of the ES-System by Ontrup et al. (2006)

browser, “Insight,” from Xerox Park. As indicated in Figure 3, the nodes of the neural network are shown as 3Dglyphs, where graphical attributes such as size or color can be used to indicate the number of documents within a node, or the document density in the semantic space, respectively. Moreover, the most frequent words are annotated to each node to give an overview of topics discussed in the information environment. The left screen of Figure 3 depicts the Poincaré-Projection of the topic’s structure. The arrow is used for navigation on the disk in a browser-like manner. In the right figure, the news activity is visualized. This tool allows the interactive exploration of the topic’s structure and the news activity in the course of time.

In a second module, the *information foraging theory* is applied to select the most relevant documents for those clusters of interest (Scholz & Wagner, 2006). In an empirical investigation, Decker et al. (2005) demonstrate this module to outperform human experts in a search task involving time-pressure. In a third module, statistical testing of association rules allows changes in previously defined topics to be monitored.

A SPOT ON FUTURE DEVELOPMENTS

Taking the user’s perspective, upcoming developments appear to be very promising. The WWW changes the traditional approach to ES in several ways. It allows smaller organizations to overcome resource barriers; it changes the role of external consultants by shifting the focus from information acquisition to consultancy; and it provides organizations with quick, up-to-date

information about competitors that was previously unavailable. Moreover, the technological development is likely to enhance ES in the business environment:

- Users’ mental models of the topics will be integrated in automated search processes.
- Even the learning progress might be integrated in software solutions customized for individual users.
- Ontology might improve users’ understanding of ES results.
- Sophisticated multimedia representation technologies (e.g., Hurson & Yang, 2005) might improve users’ interaction during the ES Process.
- The semantic Web is an up-and-coming multimedia environment for automated ES activities.

Up to now, all available software processes documents only. Other formats and technologies (e.g., MP4, Flash, or Java applets) need to be integrated. Finally, it is not only a saying that actions speak louder than words, although only a very few concepts have found their way from the academic literature to implementations that can be used for ES activities using multimedia technologies. Consequently, bridging the researcher-practitioner gap is likely to become a very profitable business.

CONCLUSION

This article describes both the domains and the approaches of ES in the business environment. The

discussion of ES activities is organized in three interpretations of ES:

- The craftsmanship-interpretation emphasizes the domains from which information fragments should be gathered and combined in order to make up a structured process.
- The art-interpretation emphasizes the human information-processing perspective of coping with different information sources, a high diversity on information quality, and, of course, differences in prior knowledge. This interpretation highlights the challenges to arrange ES in well-structured and standardized processes.
- The science-interpretation is dominated by the aim of detecting weak signals heralding relevant changes before all members of an industry become aware of the changes.

Multimedia and networks, in particular the WWW, are excellent information environments, but conventional information retrieval does not meet the requirements of the latter two interpretations. Thus, we interrelate innovative concepts of automated scanning activities to human abilities of information processing and learning to build up systems that exceed ES support for just one of the previous interpretations. Starting from a discussion of contemporarily available software prototypes the current developments in the domain are depicted briefly. This investigation provides evidence for a shift from technology-centered to human-centered approaches in developing new support tools to find, extract, understand, and use information available from electronic information environments.

REFERENCES

- Benczúr, D. (2005). Environmental scanning: How developed is information acquisition in western European companies? *Information Research*, 11(1), paper 241. Retrieved January 1, 2006, from <http://informationr.net/ir/11-1/paper241.html>
- Chen, H., Chau, M., & Zeng, D. (2002). *CI-Spider: A tool for competitive intelligence on the Web*. *Decision Support Systems*, 34, 1-17.
- Choo, C. W., Detlor, B., & Turnbull, D. (2000). *Web work—Information seeking and knowledge work on the World Wide Web*. Dordrecht: Kluwer.
- Choo, C. W., Detlor, B., & Turnbull, D. (2001). *Information management for the intelligent organization: The art of scanning the environment* (3rd ed.). Medford, NJ: Information Today.
- Day, G. S. (2002). Managing the market learning process. *Journal of Business and Industrial Marketing*, 17(4), 240-252.
- Decker, R., Wagner, R., & Scholz, S. W. (2005). Environmental scanning in marketing planning—An Internet-based approach. *Marketing Intelligence & Planning*, 23(2), 189-199.
- Garg, V. K., Walters, B. A., & Priem, R. L. (2003). Chief executive scanning emphases, environmental dynamism, and manufacturing firm performance. *Strategic Management Journal*, 24(8), 725-744.
- Hurson, A. R., & Yang, B. (2005). Multimedia content representation technologies. In M. Pagani (Ed.), *Encyclopedia of multimedia technology and networking* (pp. 687-695). Hershey, PA: Idea Group Reference.
- Jauch, L. R., & Glueck, W. F. (1988). *Business policy and strategic management* (5th ed.). New York: McGraw-Hill.
- Liu, S. (1998). Business environment scanner for senior managers: Towards active executive support with intelligent agents. *Expert Systems with Applications*, 15(2), 111-121.
- Martino, J. P. (2003). A review of selected recent advances in technological forecasting. *Technological Forecasting & Social Change*, 70(8), 719-733.
- McGoangle, J. J., & Vella, C. M. (1998). *Protecting your company against competitive intelligence*. Westport, CT: Quorum Books.
- Ontrup, J., & Ritter, H. (2006). Large-scale data exploration with the hierarchically growing hyperbolic SOM. *Neural Networks*, 19(6-7), 751-761.
- Ontrup, J., Ritter, H., Scholz, S. W., & Wagner, R. (submitted). *Environmental scanning in the WWW: Detecting important topics*.

Scholz, S. W., & Wagner, R. (2005). The quality of prior information structure in business planning—An experiment in environmental scanning. In H. Fleuren, D. den Hertog, & P. Kort (Eds.), *Operations research proceedings 2004* (pp. 238-245). Berlin: Springer.

Scholz, S. W., & Wagner, R. (2006). Autonomous environmental scanning in the World Wide Web. In B. A. Walters & Z. Tang (Eds.), *IT-enabled strategic management: Increasing returns for the organization* (pp. 212-241). Hershey, PA: Idea Group Publishing.

Tan, W. D., & Arnott, D. R. (1999). Managerial information acquisition and the World Wide Web: A theoretical background. In *Proceedings of the Fifth Conference of the International Society for Decision Support Systems*. Retrieved from citeseer.ist.psu.edu/500175.html

Walters, B. A., Jiang, J. J., & Klein, G. (2003). Strategic information and strategic decision making: The EIS/CEO interface in smaller manufacturing companies. *Information & Management*, 40(6), 487-495.

KEY TERMS

CI-Spider: Starting from a URL provided by a user, the CI Spider automatically retrieves, summarizes, and categorizes the content in linked pages using noun phrases and graphical concept maps.

Competitive Intelligence (CI): CI is characterized as (1) the process of acquiring, analyzing, and evaluating any information about known and potential competitors systematically or (2) actionable information gained in this process.

Environmental Scanning: The way in which managers study their relevant marketing environment; comprises both looking for and looking at information available in the business environment.

Information Foraging Theory (IFT): Explains information-seeking behavior of humans similar to food foraging. Basic elements are the perceived information “scent” of media (names, key words, etc.) and costs (particularly, by means of search time) for accessing the media.

Information Infrastructure: An element of IFT, which is made up by all knowledge already gained on the topic under consideration.

Information Structure: Partition of the information environment in meaningful topic clusters.

Intelligent Agents: A software entity that acts on behalf of the user to reduce both work and information load. It operates continuously and autonomously in a defined target environment.

Self-Organizing Map (SOM): A simulated neural network based on a grid of artificial neurons by means of prototype vectors. In an unsupervised training the prototype vectors are adapted to match input vectors in a training set. After completing this training the SOM provides a generalized *K*-means clustering as well as topological order of neurons.

Semantic Web: Provides us with common formats for the interchange of data related to real world objects. It is similar to the WWW, which enables the interchange of documents, but aims to become an environment for automated processing in addition to human browsing.