

VOLUME 10 (2006): ISSUE 1. PAPER 0**What does the Web represent? From virtual
ethnography to web indicators****Andrea Scharnhorst*, Peter van den
Besselaar**,***, Paul Wouters****The Virtual Knowledge Studio. KNAW. Amsterdam
The NetherlandsE-mail: andrea.scharnhorst@vks.knaw.nlE-mail: paul.wouters@vks.knaw.nl

** Amsterdam School of Communications Research, University of Amsterdam

*** Rathenau Institute (Science and Technology Assessment)

E-mail: P.vandenbesselaar@rathenau.nl

The Web can be seen as a place where scholarly and other activities become visible in a specific way. Web analyses vary from the structural and statistical analysis of huge web graphs to virtual ethnography. Any use of web data is linked with expectations. One expects to see something that is not visible in the off-line world, that represents a new quality, or that indicates certain novel interactions.

The papers in this special issue explore the perspective from which it is possible to link quantitatively oriented web data collection to a qualitatively informed one. Research questions addressed by this combination of methods reach from the usability and newsworthiness of hyperlinks for different audiences, to disciplinary differences in web use, to the relevance of web indicators in science policy context.

Both quantitative and qualitative studies struggle with the meaning of the hyperlink, with the problem of how to draw boundaries around the object of study, and with the definition of the best unit of analysis. In the interpretation of web data, the researcher is moreover faced with the necessity of reflecting upon the nature of the Web as a set of converging media and the relation between off-line and on-line worlds.

This special issue encompasses papers that analyze web data gathered with various methods for various purposes. The goal of the issue is to explicate the assumptions under which web data gathering takes place. The authors make visible how initial expectations shift during the process of data collection and interpretation.

We draw specific attention to the need and the potential of the combination of qualitative and quantitative analyses. This combination, which is of course not specific to Web analysis, can be realized in different ways. Huge data samples can be collected at different levels of the science systems, and their interpretation can be linked to traditional questions in science and technology studies about the formation, structure and function of specialties and regularities in the knowledge production. Automatic retrieval of web data from search engines can be combined with an interview based qualitative approach, both to understand the function of hyperlinks in communication and to draw boundaries around the unit of analysis. Quantitative approaches as network analysis can also be used to trigger ethnographic observations that in turn deliver a rich spectrum of possible meanings and functions of hyperlinks. To encourage a debate between the different approaches we have chosen a special format in this issue. After each paper a discussant comments on the

The Journal

Cybermetrics News

Editorial Board

Guide for Authors

Issues Contents

**The Seminars****The Source**

Scientometrics



Tools



R&D Policy & Resources



most important and interesting points and explores possible extensions of the research.

The issue starts with a paper by Gaston Heimeriks and Peter van den Besselaar (Analyzing hyperlinks networks: The meaning of hyperlink based indicators of knowledge production) which gives an overview of the creation of web indicators, concentrating on hyperlink networks at various levels (countries, universities, research departments, individual research). They find that language and geographic aspects govern the structure of hyperlink networks at the level of countries. It might be at the level of research departments that we can expect to see integrating tendencies in the European Research Area.

If one wishes to study the communication behavior within different specialties on the web, sociological approaches to science studies can be used as a framework to interpret apparent differences. Jenny Fry (Studying the scholarly web: how disciplinary culture shapes online representations) develops methods of a qualitative hyperlink analysis. Combining interview data, crawler results and search engine data, she uses the concept of embedded webspheres by applying and further developing the websphere proposed by Foot and Schneider.

While in the case of Jenny Fry's analysis the unit of analysis (websphere) is constructed in a semi-quantitative way, in the paper by Viv Cothey, Isidro Aguillo and Natalia Arroyo (Operationalising "Websites": lexically, semantically or topologically?) a web site is defined using automatic link harvesting. Different ways to extract information from web graphs: using the lexical structure of an URL address, the semantic analysis of the topics in a web page, and the analysis of the topological structure of the web graph are discussed. This last approach as proposed and tested by the authors, opens up the possibility to identify hidden community structures on the web using topological clusters as a form of "web spectroscopy". It also points to the need for a qualitative interpretation of the structures found on the Web in terms of the cognitive or institutional organization of science.

One way to provide such meaning is to use smaller case studies. In the paper of Eleftheria Vasileiadou and Peter van den Besselaar (Linking shallow, linking deep. How scientific intermediaries use the web for their network of collaborators), the authors start with an institutionally defined web site to look into a certain aspect of the structure of the out-link network (depth of a link) to map and analyze web visibility of collaborations. Such an analysis provides insights in the use of the web technology by institutions and thereby provides a context for larger hyperlink studies.

The rich symbolic nature of hyperlinks with different functionalities becomes especially visible if virtual ethnography is applied. Anne Beaulieu and Elena Simakova (Textured connectivity: an ethnographic approach to understanding the timescape of hyperlinks) combine in-link and out-link analysis with link context. Looking at the web presence of different databases they explore different ways of referencing around hyperlinks. In this way the networked character of knowledge production in e-science becomes visualized. If the context of the hyperlink is taken into account, a temporal history of the diffusion of information on the web can be written.

The request in the first paper of this special issue for longitudinal web analysis is emphasized also in the last article of this special issue. Andrea Scharnhorst and Paul Wouters (Webindicators – a new generation of S&T indicators) discuss the possible role of web indicators in science policy contexts as process indicators. The paper describes the conceptual basis for the WISER project. They point to possible intersections between complex network research and webometrics in the development of web indicator reflecting the dynamic and multi-faced character of hyperlink networks.

To sum up, with this special issue, we hope to contribute to the development of on-line methodologies in science studies, the conceptualization of the web, and the understanding of the role of the web for science and technology.

Finally, we would like to mention that the papers in this special issue have been presented at a special session at the conference "Public Proofs: Science, Technology and Democracy" 4S-EASST Conference 2004, held in Paris, 25-28

August 2004. To document and continue the interesting discussions we had around this session we decided to introduce a discussant to each paper.

Part of this work was supported by a grant from the Common Basis for Science, Technology and Innovation Indicators part of the Improving Human Research Potential specific programme of the Fifth Framework for Research and Technological Development of the European Commission. It is part of the WISER project (Web Indicators for Scientific, Technological and Innovation Research) (Contract HPV2-CT-2002-00015) (www.webindicators.org). Part of the work has been funded by the Fifth Framework Program of R&D of the European Commission within Cross Program Action CPA4 of the Information Society Technologies Programme (IST) by the European Commission, grant IST-1999-20350, the EICSTES project (European Indicators, Cyberspace and the Science-Technology-Economy System) (www.eicstes.org).

Table of content

Gaston Heimeriks and Peter Van den Besselaar
What do hyperlinks mean? The value of hyperlink networks as indicators of knowledge production
Discussant: [Viv Cothey](#)

Jenny Fry
Studying the scholarly web: how disciplinary culture shapes online representations
Discussant: [Eleftheria Vasileiadou](#)

Viv Cothey, Isidro Aguillo and Natalia Arroyo
Operationalising "Websites": lexically, semantically or topologically?
Discussant: [Gaston Heimeriks](#)

Eleftheria Vasileiadou & Peter van den Besselaar
Linking shallow, linking deep: how scientific intermediaries use the web for their network of collaborators
Discussant: [Anne Beaulieu](#)

Anne Beaulieu, Elena Simakova
Textured Connectivity: an ethnographic approach to understanding the timescape of hyperlinks
Discussant: [Jenny Fry](#)

Andrea Scharnhorst, Paul Wouters
Webindicators – a new generation of S&T indicators
Discussant: [Peter van den Besselaar](#)

Received 15/Dec/2005

Accepted 31/Jan/2006