

February 1999

Academic Journals on the Internet

Alexander Hars

University of Southern California, Los Angeles, USA, Marshall School of Business, hars@bus.usc.edu

Follow this and additional works at: <http://aisel.aisnet.org/wi1999>

Recommended Citation

Hars, Alexander, "Academic Journals on the Internet" (1999). *Wirtschaftsinformatik Proceedings 1999*. 6.
<http://aisel.aisnet.org/wi1999/6>

This material is brought to you by the Wirtschaftsinformatik at AIS Electronic Library (AISEL). It has been accepted for inclusion in Wirtschaftsinformatik Proceedings 1999 by an authorized administrator of AIS Electronic Library (AISEL). For more information, please contact elibrary@aisnet.org.

Academic Journals on the Internet

Alexander Hars

Marshall School of Business USC, Los Angeles, USA (hars@bus.usc.edu)

Contents

- 1 Introduction**
- 2 Traditional electronic journals**
 - 2.1 Hybrid electronic journals
 - 2.2 Journal of Artificial Intelligence Research (JAIR)
 - 2.3 MISQ Discovery
 - 2.4 Evaluation of characteristics
- 3 Encyclopedia / Knowledge Repository**
 - 3.1 GenBank
 - 3.2 Foldoc
 - 3.3 Evaluation of characteristics
- 4 Collaboratory**
 - 4.1 Upper Atmospheric Collaboratory (UARC)
 - 4.2 Information Systems Cybrarium
 - 4.3 Evaluation of characteristics
- 5 Other types of sites**
 - 5.1 Preprint archive
 - 5.2 Ontolingua Server
 - 5.3 JSTOR
 - 5.4 Evaluation of characteristics
- 6 Synthesis: Characteristics of electronic journals**

Abstract

While much research examines the internet's impact on 'electronic business' information technology may have an equally profound impact on what might be called 'electronic science'. Although the internet has the potential of revolutionizing research, it is not clear in which ways the structure of journals and the associated research processes will be affected. The purpose of this article is to provide a better foundation for the engineering of scientific knowledge infrastructures on the internet. Alternative models of academic sites which are dedicated to the creation and dissemination of scientific knowledge are studied. Characteristics of different categories of electronic journals, e.g. traditional electronic journals, knowledge repositories, collaborations etc. are examined and synthesized into a list of key characteristics which then can be used as basis for the design and investigation into the evolution of such systems.

1 Introduction

The internet challenges traditional models of scholarly communication. Although it is possible to replicate the structure of current journals on the world wide web, many researchers argue that research processes need to be fundamentally rethought:

"The new electronic medium gives us the opportunity to reconsider many aspects of our current research communication, and researchers should take advantage of this opportunity to map out the ideal research communication medium of the future." ... "In particular, we need to dislodge definitively the curiously prevalent notion that the future electronic medium will strictly duplicate, inadequacy for inadequacy, the current print medium." (Ginsparg, 1996)

The new medium presents several opportunities for reinventing scholarly communication. On the internet, the costs of publishing are dramatically reduced. Some authors (e.g. Odlyzko 1997) argue that they are practically zero. This has significant implications for the role of reviewed journals and scholarly associations. In contrast to paper, electronic knowledge can be modified easily. Links, indexes and meta structures can be created on existing knowledge. The presentation of knowledge can be tuned to the needs of the reader and there may be many additional innovations which are enabled by information technology.

The new medium therefore invites the reengineering of scientific processes. Many scientific associations, research institutes and individual researchers have begun to leverage information technology for the creation and dissemination of scientific knowledge. In most of the cases, however, familiar journal structures are being replicated (Hars & Malhotra, 1997). A recent example is the

introduction of two new electronic journals by the Association of Information Systems (AIS). Adherence to traditional structures may have several reasons: 1) there is uncertainty about alternative models and 2) a lack of experience about their feasibility. 3) In addition, current research processes and the reward system have been optimized for the traditional structure. Therefore the replication of these structures may be the least risky strategy.

The purpose of this article is to provide a better foundation for the engineering of scientific processes on the Internet. Alternative models of academic sites which are dedicated to the creation and dissemination of scientific knowledge will be studied. While it may be too early to determine whether the evolution of electronic journals will converge to any of these prototypes, this research will show the breadth of change that is possible in comparison to the traditional journal model.

The article is structured as follows: To identify the spectrum of possible knowledge-related sites, characteristics of current sites are identified. Characteristics are elicited by comparing electronic sites with traditional media and by comparing different electronic sites. The characteristics are then prioritized and organized into groups of key characteristics. This allows to extrapolate the full spectrum of possible knowledge related sites by considering all combinations of independent characteristics. In the following, sites will be examined by categories typically mentioned in the literature such as: electronic journal, knowledge repository, collaborative. Then sites will be examined which do not easily fit into these categories.

2 Traditional electronic journals

Traditional electronic journals are scholarly sites on the World Wide Web which closely replicate the structure of paper journals. The qualifier 'traditional' is used to distinguish the concept from the broader vision of electronic journals as "scholarly IT based infrastructures" (Watson 1994) which will eventually replace paper journals. Many traditional electronic journals now exist on the WWW.

2.1 Hybrid electronic journals

Most publishers of paper journals have begun to make online versions available. Examples are the Association for Computing Machinery, IEEE, Elsevier and Springer. These sites are hybrid sites; journals are published both on the internet and in paper. In most cases access is restricted to subscribers of the electronic journal (or an electronic library); ACM also provides pay-per-view access to individual articles. The presentation of the journals on the web resembles the traditional structure. The electronic journals are organized by volume and issues: tables of contents provide access to individual articles. In most cases, access to abstracts

and keywords is free. Articles typically have an index page which shows title, abstract and keywords and provides a link to the full text. Full text is mostly available in PDF or Postscript format. Full text in html is rare. As an alternative form of accessing articles, most sites provide a search engine. Search technology is advancing rapidly and the range of capabilities varies considerably. Most search engines provide full-text search; some provide context-based search or searches by article attributes such as publishing date, keywords etc.

Hybrid sites are necessarily very close to paper journals. Most elements of the publishing process remain essentially unchanged. In the case of the above mentioned journals, for example, accepted papers can not be changed. The review process remains essentially the same (peer review) although some of the journals now accept online submissions and some of the sites are moving toward online support for the review process.

2.2 Journal of Artificial Intelligence Research (JAIR)

Interesting innovations have been provided by some electronic journals. The Journal of Artificial Intelligence Research, founded in 1993, describes itself as one of the first electronic journals. It is a hybrid journal because each of its volumes is published in paper form. However, no issues are published. Similar to traditional journals, the table of contents is organized by volume and author. Recently an innovative subject area index has been added: readers can browse a JAVA applet which graphically clusters articles by subject area. Another innovation is that articles are published online as soon as they are accepted - rather than to wait for the official publishing date of the journal. This approach is taken by an increasing number of traditional electronic journals, including ACM and MISQ.

2.3 MISQ Discovery

MISQ Discovery is a venture of MISQ, one of the leading information systems research journals. It was announced in 1994 as "an adventuresome and experimental electronic production" (Ives, 1994) which, among others was intended to "challenge the foundations of paper-based scholarship" (Ives, 1994). The first article was published in September 1996 (Ytterstad, 1996). Although very few articles have been published so far, the journal has introduced some innovations: It pioneered the distinction of living vs. archival articles. Archival articles can only be subsequently edited to correct typing, spelling, grammar, or to update links (<http://www.misq.org/discovery/about.html#living>, 3/28/1998). In contrast, living publications are kept up-to date by the authors. In addition, for the first article a listserver-based bulletin board was added to allow discussion of the papers. This feature was discontinued, however, because of apparent lack of interest. MISQ Discovery uses a traditional peer-review system. To improve the

acceptance of the journal, an abstract of each accepted article is published in MISQ and the article becomes part of the MISQ index.

2.4 Evaluation of characteristics

The preceding analysis of traditional electronic journals has shown several characteristics which need to be distinguished: hybrid* vs. electronic-only site; fee structure (free, subscription*, pay-per view) organization of knowledge (vol./issue*, subject tree/diagram), search mechanisms, format for articles (PDF*, PS*, HTML), online submission*, online review*, pre-publishing of accepted articles, archival/static* vs. living/updateable articles.

The degree of innovation vs. replication of traditional structures can be analyzed in each of these characteristics. A * indicates close correspondence to the traditional model. Online submission and online review as well as PDF and PostScript formats are included in this category because they replicate the traditional process in the electronic environment rather than being fundamentally rethought.

A characteristic which merits a closer look is the practice of pre-publishing accepted articles. This practice indicates a breakdown of the traditional paradigm. Traditional electronic journals have introduced this feature, because it became obvious that there is no need to wait for publishing an accepted article until all articles have been reviewed. This does not address the fundamental issue, however: the obsolescence of the issue metaphor. The economics of paper prohibit publishing individual articles - therefore issues were needed. In the electronic environment, this need does not persist. The characteristic is not an innovation but a work-around. As a consequence, it will not persist in mature electronic journals which will cease to be organized by issue.

3 Encyclopedia / Knowledge Repository

An important alternative to paper journals are encyclopedias - "works that claim to provide in an orderly arrangement the essence of "all that is known" on a subject or a group of subjects" (Britannica Online 1998 <http://www.eb.com:180/cgi-bin/g?DocF=macro/5002/11/0.html> (9/14/1998)). On the World Wide Web, several scientific sites with such a goal have emerged. To differentiate such sites from traditional electronic journals we propose the term 'knowledge repository'. This term alludes to the goal of storing knowledge in a large, dynamic and well-structured knowledge base.

Examples of such sites are:

3.1 GenBank

GenBank is an electronic database maintained by the National Center for Biotechnology Information (US). It contains more than one billion annotated and cross-referenced DNA sequences. Data is regularly exchanged with other DNA databases located in Europe and Japan. Scientists can submit their DNA sequences to GenBank free of charge on the web. Before a submitted sequence is published, it is reviewed and indexed by GenBank staff. Not all GenBank submissions are accessible by the public. On demand by the author, access to a sequence can be blocked (GenBank, <http://www.ncbi.nlm.nih.gov/Web/Newsltr/feb98.html> #GenBank, 6/23/98). GenBank incorporates a revision mechanism to allow authors to update their sequences. An important functionality for gene researchers is the ability to perform similarity searches between DNA and protein sequences using software which was specifically developed for this purpose. Many scientific journals require a submission to GenBank as a prerequisite for publishing an article. While GenBank was not intended as a competitor to academic journals, it has improved the ability of the research community for gathering high-quality data. Although no systematic studies have been executed yet, it may have shifted the focus of research efforts more towards data. Cinkosky et al. (1991) concluded: "Our experience indicates that electronic data publishing has actually resulted in a higher level of quality for data in GenBank and therefore also in journals".

3.2 Foldoc

Another example of a knowledge repository is the Free Online Dictionary of Computing (Foldoc) (<http://wombat.doc.ic.ac.uk>). It has been available on the World-Wide Web since 1994 and contains more than ten thousand short but intensively cross-referenced entries. The dictionary has a single editor who maintains tight control of its content; contributions of other authors are solicited and a mechanism for guest editors has been set up, but entries do not list the author. A main page, however, lists more than 800 individual contributors. An interesting feature of the dictionary is a list of missing entries which is based on a log of reader's unsuccessful searches.

3.3 Evaluation of characteristics

Most of the characteristics identified for traditional electronic journals also apply for encyclopedias. One main difference is the emphasis on the organization of knowledge which leads to an organization by subject area rather than an organization by volume/issue. Another characteristic are algorithms for the detection of similarities between different units of knowledge. They can be viewed as specific types of search algorithms. Another characteristic is the ability to add private contributions which can only be viewed by the author. In contrast to most

of the traditional electronic journals, which use static hypertext, knowledge repositories often dynamically generate html pages.

Thus the following additional characteristics may be relevant: technology (static hypertext* vs. dynamic (database)), absence* vs. existence of different levels of access restrictions to articles, absence* vs. existence of similarity search algorithms,.

4 Collaboratory

The origins of the term collaboratory can be traced back to a 1989 NSF workshop (Lederberg & Uncapher 1989; Finholt & Olson 1997). In 1993 the National Research Council defined the term collaboratory as: "A center without walls, in which the nation's researchers can perform their research without regard to geographical locations - interacting with colleagues, accessing instrumentation, sharing data and computational resources [and] accessing information in digital libraries" (National Research Council 1993, p. 7). Barua et al. (1995) defined it as "an open electronic platform for individuals or groups with common interests to efficiently exchange, disseminate and create issues, ideas and knowledge" (1995, p. 418).

4.1 Upper Atmospheric Collaboratory (UARC)

After the need for collaboratories had been identified space physics was one of the first areas where the concept was implemented. With funds from the National Science Foundation, the Upper Atmospheric Collaboratory was started in 1992 at the University of Michigan in collaboration with several other research institutes around the world. The first step consisted in tying together instruments operated in various sites and allowing uniform access to the stream of data which they generate. The infrastructure subsequently was extended to include chat capabilities and Java-based applets for the visualization of data. In addition, access to computer models was provided. In addition, UARC provides the ability to conduct 'research campaigns' during which researchers from several institutes across the world and usually from several disciplines pool their resources to perform joint research. Based on a coordinated research plan, a campaign allows online discussion of data as it becomes available and adjustment of instruments. Part of the data is archived via UARC and accessible for subsequent analysis.

4.2 Information Systems Cybrarium

The [Information Systems Cybrarium](#) is a web-based collaboratory which has been developed at the University of Southern California by the author. Its goal is to investigate new ways for the creation, synthesis and application of information

systems knowledge. The current Cybrarium prototype is used to explore different alternatives on the path towards more advanced web-based knowledge infrastructure and to increase our understanding of the impact of information technology on research processes. While many current academic internet sites related to IS knowledge are modeled on traditional journals and thereby are focused on the dissemination of knowledge, the Cybrarium is focused on the whole life-cycle of knowledge from creation through synthesis to dissemination.

Information is stored in the Cybrarium as a collection of typed nodes and typed relationships between nodes. Nodes and relationships are stored in a database which is linked to the Web. The use of a database has the advantage that the presentation of the knowledge on the web can be separated from the storage of knowledge. It becomes possible to assemble knowledge for the reader from parts of the network of nodes and relationships. In contrast to the World Wide Web, where links can only be traversed in one direction, the utilization of a database also has the advantage, that links can be traversed in both directions. Similar to the web, nodes can represent multimedia information.

The Cybrarium currently is accessible at <http://cybrarium.usc.edu>. A screen shot of the main menu is shown in figure 1. As this is a collaborative infrastructure, users are encouraged to identify themselves through registration. Consistent with the node & relationship approach a node is created for each registered user. All submissions, evaluations etc. are subsequently linked to that node. Registration is not required, however, for read access. Currently email feedback is used to authenticate the user. With approximately 1200 nodes, the knowledge content of the Cybrarium is small but its growth rate has accelerated over recent months.

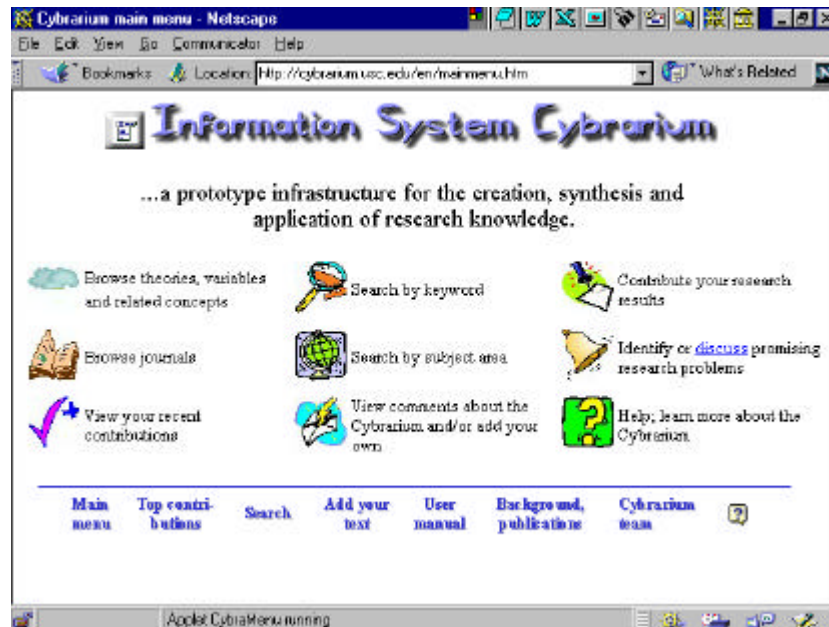


Figure 1: Information Systems Cybrarium

In the following, some of the key features of the Cybrarium are briefly discussed:

Instant publishing: All contributions to the Cybrarium are instantly published on submission without review. This is not a disadvantage for the reader as she has the ability to suppress un-reviewed articles through a filter (see below on collaborative filtering). Instant publishing eliminates publishing delays and increases the collaborative nature of the system.

Collaborative filtering: Each reader is encouraged to rate a paper along three dimensions: relevance, rigor, and level of detail. It takes only three mouse clicks to assign the 3 ratings. The ratings can be used to set a filter. A practitioner, for example, may elect to view articles which are highly relevant - independent of the rigor rating while a PhD student may elect to view primarily papers with high levels of rigor etc.

Annotation & notification: Every reader can add comments to an article. Readers are automatically informed via email if a comment is posted.

Open architecture: Contributions to the Cybrarium need not be physically stored on the Cybrarium site. Articles may be stored anywhere on the world wide web. When an article is requested that is stored externally, the Cybrarium displays its content in a frame. A second frame provides access to Cybrarium functionality such as annotation, related topics, ratings etc. for the article. The Cybrarium thus adds a layer of hypertext functionality on top of the World Wide Web.

Meta journal: New articles appearing in the top journals in Information Systems are systematically incorporated in the Cybrarium - either via an external link to an abstract on the journal site or via an abstract. Currently 8 Ph.D. students and faculty are working on this task of adding 'seed knowledge' to the Cybrarium.

Identification of research problems: Traditional journals focus on the result of the research process - the dissemination of knowledge. The Cybrarium focuses on the whole process from creation to synthesis and dissemination. An example is a functionality which allows researchers to identify and discuss promising research problems in an area.

4.3 Evaluation of characteristics

Collaboratories add a wealth of additional characteristics: conditional* vs. instant publication of contributions, absence* vs. existence of custom filters, closed* vs. open knowledge infrastructure, unidirectional* vs. bidirectional cross-references, absence vs. existence of reader-based review/rating of articles.

5 Other types of sites

5.1 Preprint archive

Another variant of web-based academic journals are preprint archives. The first preprint archive was started by Ginsparg to simplify the distribution of preprints among high energy nuclear physicists (Ginsparg 1994). The site has expanded to more than 30 subject areas including sections for economics and mathematics and now receives more than 400.000 requests per week (cf. http://xxx.lanl.gov/cgi-bin/show_weekly_graph (9/12/1998)). Preprints are submitted to one of the subject areas and instantly published without review. Mechanisms for cross-listing a preprint in several sections exists. Readers are presented with daily lists of new submissions which show title, authors and abstract for each preprint and which contain a link to the text file. The system has search capabilities. The archive supports multiple versions of a document. It is possible to change preprints after submission; new versions receive a new identifier. However, preprints can not be linked to one another. This has led to first changes for academic journals in physics: Several scientific physics journals have started to accept preprints from the archive for submission. To submit a paper to one of the American Physical Society Journals (<http://publish.aps.org/esub/xxxsub.html>) or to the Journal of High Energy Physics, for example, an author can electronically submit the identifier of his preprint (<http://jhep.cern.ch/JOURNAL/xxx.html>). The editor, then downloads the document from the archive. The concept has been applied in other areas.

Preprint archives have been established, for example, for engineering (<http://www.eevl.ac.uk>) and mathematics (<http://www.ams.org>).

5.2 Ontolingua Server

[Ontolingua server](#) is a collaborative modeling environment which allows the development and reuse of ontologies. Ontologies are formal knowledge representations for specific subject areas. Ontolingua server provides a customized ontology editor for the creation and maintenance of ontologies. It contains analysis routines to check the consistency of ontologies and allows incorporating existing ontologies such as an ontology about space and time. This reduces the effort of building an ontology. Ontolingua server is an interesting system because it provides a custom interface for the specification of (admittedly very formal) knowledge. It identifies dependencies between different units of knowledge and allows reusing knowledge created by others - a simple form of collaboration.

5.3 JSTOR

JSTOR is an archive of digitized (bitmapped) articles from print journals which are older than three years. More recent articles have been excluded to minimize effect on the revenue stream of publishers while improving access to older articles. JSTOR was founded in 1995 with initial funding from the Mellon foundation. In 1998, JSTOR listed more than 300 participant libraries and is nearing its goal of providing access to the content of 100 journals in 15 disciplines.

5.4 Evaluation of characteristics

This review has led to the identification of the following additional characteristics: Absence* vs. existence of consistency checking algorithms; Absence* vs. Existence of possibility to reuse/integrate other units of knowledge, Absence* vs. existence of versioning mechanisms.

6 Synthesis: Characteristics of electronic journals

The previous analysis has led to the identification of a significant number of characteristics of electronic journals. They can be aggregated into five main groups which are shown in Table 1. The first group contains characteristics which refer to the mode of interaction. While traditional electronic journals focus primarily on read/browse, collaboratories add the dimension of authoring and evaluation. Additional characteristics are the existence or absence of online annotation, review and submission of texts. The second group refers to the organization of knowledge. Key issues are the organizing principles: The traditional journal is organized temporally by volume and issue. Other sites may

provide category networks such as subject trees etc. or provide active mechanisms to ensure the consistency of contributions (e.g. Ontolingua server). Another characteristic is the number and type of indexes provided (e.g. by author, by subject area, by institution, by research problem, by theory etc.). 'Link semantics' refers to the types of relationships and cross-references which can be distinguished. On the world wide web and in the print medium, for example, references are unidirectional. Two additional characteristics in this group are consistency mechanisms and versioning. The third group refers to the presentation of knowledge. Navigation aids (e.g. knowledge maps) and filtering mechanisms which allow a reader to customize the journal need to be distinguished. User models and internationalization are issues which are currently not present in the electronic journals reviewed above; they may become increasingly important in the future. The fourth group refers to the underlying technology. Many advances are currently being made in search technology; a distinction needs to be made between journals which use static hypertext and those which rely on database technology. Only Ontolingua server has an inference engine. Additional features in this group are the utilization of a state protocol which allows sessions between reader and electronic journals and support for multi-media. The final group relates to governance. The target audience of a journal, revenue model (e.g. subscription, pay-per view, free) as well as authentication mechanism need to be distinguished. Additional critical issues are review and reward mechanisms.

Interaction	Organization of knowledge	Presentation	Technology	Governance
Interaction types	Organizing principle	Navigation aids	Search engine	Audience
Online annotation	Indexes	Filtering mechanisms	Database	Revenue model
Online review	Link semantics	Internationalization	Inference engine	Authentication
Electronic submission	Consistency mechanisms	User models	State/session protocol	Reward mechanisms
Email notification	Versioning		Media types	Review mechanisms

Table 1: Characteristics of electronic journals

These characteristics can be utilized as guideline for the design of electronic journals. The table shows that the variety of possible approaches is very high. Electronic journals of the future will not replicate traditional paper journals. As a consequence, scientific processes and reward systems are likely to be changed. It is time to experiment with better alternatives for scientific knowledge management infrastructures and to reengineer traditional research processes.

References

- Barua, A./Chellappa, R./Whinston, A.B (1995). Creating a collaboratory in cyberspace: Theoretical foundation and implementation. In: *Journal of Organizational Computing*, (5:4), pp. 417-442.
- Britannica Online (1998): Version 98.2. <http://www.eb.com> (9/14/1998).
- Cinkosky, J./Fickett, J.W./Gilna, P./Burks, C. (1991): Electronic data publishing and GenBank. *Science* 252: 1273-1277.
- Finholt, T.A./Olson, G.M. (1997): From laboratories to collaboratories - a new organizational form for scientific collaboration. *Psychological Science*, (8:1), pp. 28-36.
- Ginsparg, P. (1996): Winners and losers in the global research village. Presented at UNESCO conference, Electronic publishing in science, February, 1996, <http://xxx.lanl.gov/blurb/pg96unesco.html> (4/26/1998).
- Hars, A./Malhotra, A. (1997): Emerging web-based knowledge bases: Categories, trends and implications for the IS research community. <http://cybrarium.usc.edu/cyb.dll/?object=1000610>.
- Ives, B. (1994): MISQ Central: Creating a new intellectual infrastructure. Editors Comments, *MISQ* Vol. 18, No. 3, pp. xxxv-xxxix.
- Lederberg, J./Uncapher, K. (1989): Towards a national collaboratory: Report of an invitational workshop at the Rockefeller University. Washington, D.C.: NSF, Directorate for Computer and Information Science and Engineering.
- National Research Council (1993): National collaboratories - Applying information technology for scientific research. Washington (National Academy Press).
- Odlyzko, A. (1997): The Economics of Electronic Journals. Proceedings of the conference on Scholarly Communication and Technologies. <http://arl.cni.org/scomm/scat/odlyzko.html> (8/20/1998)
- Watson, R. (1994): Creating and sustaining a global community of scholars. *MISQ*, Vol. 18, No. 3, 228-231. <http://www.misq.org/archivist/vol/no18/issue3/vol18n3art1watson.html>
- Ytterstad, P./Akselsen, S./Svendsen, G./Watson, R.T. (1996): Teledemocracy: Using Information Technology to Enhance Political Work. *MISQ Discovery*, 9/1996, <http://www.misq.org/discovery/articles96/article1/> (9/10/1998).