Tilburg University

A user's view on the electronic library

van Groenendaal, W.J.H.

Published in:
Library Acquisition: Practice and Theory

Publication date:
1997

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
van Groenendaal, W. J. H. (1997). A user's view on the electronic library. *Library Acquisition: Practice and Theory*, 21(3), 337-345.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Pergamon

Library Acquisitions: Practice & Theory, Vol. 21, No. 3, pp. 337-345, 1997
Copyright © 1997 Elsevier Science Ltd
Printed in the USA. All rights reserved
0364-6408/97 \$17.00 + .00

PII S0364-6408(97)00062-8

**THE FIRST ELSEVIER ELECTRONIC SUBSCRIPTIONS CONFERENCE
OCTOBER 17-18, 1996
HEEMSKERK, THE NETHERLANDS**

A USER'S VIEW ON THE ELECTRONIC LIBRARY

WILLEM J. H. VAN GROENENDAAL

Lecturer, Department of Information Systems and Auditing
Center for Economic Research (CentER)
Faculty of Management and Economics
Tilburg University
5000 LE Tilburg
The Netherlands
Internet: w.j.h.vgroenendaal@kub.nl

Abstract—*The electronic library offers researchers the opportunity to consult more information sources than ever before. The times of long visits to the library to browse through books and journals belong in the past for many of us. In addition, the personal relationships with experienced library staff who knew your interests has disappeared. This has been replaced by search profiles that supply you with regular updates on new publications in your field of research. The electronic library is, however, more than a digitalized version of the old library. Searching has become easy and quick, and the number of databases that can be accessed has increased significantly. For most users the library is, however, no longer the only, or even the main, road along which information becomes available. It is used in combination with the many sources that have emerged along the electronic highway. The main problem for the user is how to manage all these different resources. Although the library information system is relatively easy to use and shows a large number of user-friendly extras, it is not what he/she really needs. What the user really needs is intelligent assistance in using this combination of local and remote information sources. However, nothing is for free. The question is how are we going to pay for all these new opportunities in a time of diminishing budgets. The answer may be to make full use of the electronic library.* © 1997 Elsevier Science Ltd

Keywords—Electronic library, Collection usage, Specialized databases, Cost reduction, Intelligent assistance

INTRODUCTION

The Internet has changed the way we can exchange information. Some even speak of a revolution that will have the same impact as the invention of printing and therefore call it the PostGutenberg Galaxy [1]. The aim of this paper is not to discuss all new opportunities of the Internet in combination with electronic information storage; see the paper by Furner-Hines and Wilett [2] for this. This article contains a user's view on what an up-to-date academic library using modern information technology (IT) should offer and how it can affect the work of scholars and scientists. Some wishes for future support of scientific research are formulated. Note that the electronic library here is restricted to academic libraries that contain the output of and facilitate scientific research.

The author is an interested user only, not a librarian. An interested user, because: (1) I head the faculty committee that allocates the budget on books, journals, and databases over the different subjects of interest to the faculty in my area; (2) at the same time I am responsible for the budget of my department; and (3) last but not least I spend almost half of my time doing research in the field of decision support and simulation. In the typology of Pagell [3] I qualify as a knowledgeable end user. My view on the electronic library is not based on research in the field of electronic information exchange, but on using the library of Tilburg University in The Netherlands, which is one of the most modern libraries in the world.

Until the end of 1991 we had a separate traditional library for each faculty, and one central library containing books, statistics, and journals of general interest. In 1992 the university opened a completely new centralized library, equipped with the most advanced fiberoptics-based network that money could buy, connecting the central computers, all computers used by the staff, and approximately 450 personal computers within the library for public use. Everybody was equipped with an "integrated workplace," that is, has access to the library system, the computer center, a number of databases, the World Wide Web (WWW), and some other facilities. As a result, much of the communication between staff and students is now based on our local network and the Internet instead of personal contact. At the same time, the library and Elsevier Science started to develop library information systems that would facilitate the use of electronic journals.

For the scientific staff, all these new facilities opened up opportunities never dreamt of before. With the connection to the Internet, the library is no longer the only source of information, but one of the many resources available on the WWW, where a number of interesting experiments to exchange information are occurring. The most famous example of how the WWW can affect scientific information exchange are the E-Archives in Los Alamos. The E-Archives have become the core of information for researchers in the field of high energy physics and related topics [4] and are enlarged daily and visited by physicists from all over the world (up to 50,000 connections per day from more than 80 countries). Some examples in my own field of interest are:

- CORRYFEE, a site plus mailing list edited by Fethi Mili, through which new research papers in economics and management from all over the world are announced [5];
- the Economics Research Network (ERN) run by Martin Feldstein of Harvard University and Micheal C. Jensen of Harvard Business School [6] contains specialized lists covering all areas of economic research;
- SNET DIGEST, a mailing list edited by Shane Henderson on events in the field of simulation; and
- the SAMO site, where scholars from various disciplines exchange information on sensitivity analysis of model output.

Web sites all over the world offer information that can be downloaded and printed. Many use hypertext, which means that you can get not only a discussion paper, but also related publications or even the computer program used to obtain the results. As a result, the old ways to gather information seem to wear out. Although most researchers still have private subscriptions to the most important journals in their field (the author has six), this source of information is becoming less and less important. The paper on which a journal article is based is now made available much earlier through the WWW (be it in its pre-refereed form). Such an early version of new ideas is needed because of the long delays between the moment an article is submitted and the moment it is published (often up to 3 years). In combination with the increased pressure on scientists to publish, the new information sources play an important role in research activities. However, the traditional paper stream of information is still in place too. Or worse, the stream sent by the suppliers of books, periodicals, and databases still seems to grow. The Internet has led to an exponential increase in the amount of electronically available information (by no means all of it is useful), which exists alongside the printed stream. As a result, the interested user can easily drown; many long for the good old days. Several of my less computer literate colleagues lost track, and do not or only partly use the new opportunities.

This paper will address several questions. What are the advantages of electronic information resources? Next, we discuss new opportunities the electronic library in combination with the WWW can offer scientists. Will the Internet change the role of the library from information storekeeper to information navigator? What are the expected costs and benefits? The paper will end with some conclusions.

NEW SYSTEM, NEW RULES

Many of my colleagues do not (yet) fully appreciate the new library system; some even say they long for the good old days. What were these good old days? In the past, every faculty and some departments had their "own" library close to the office, containing mainly books and journals. If you needed anything you went down a few stairs to collect it. One of the main advantages was that you could browse through journals and books, which often resulted in information on other subjects than the one you were initially seeking. The library personnel knew the scientific staff personally, and if you needed anything special you had easy access to their knowledge. Only occasionally and after you were granted a budget did you use a computerized literature search in one of the large databases in your field of research (mostly located in the U.S.). Furthermore, since the library staff knew your research interests, they would often draw your attention to new books or journals. Also, the strict rules for not lending out journals, or not providing information on who had borrowed the book you urgently needed, could easily be bent in your favor. It is the browsing and the personal contact that many of my colleagues miss the most.

In the new centralized library the human and personal approach is, to a large extent, replaced by assistance based on software. The library's hardware is designed to meet the IT possibilities of the future. To fully utilize these IT possibilities, the job descriptions of library staff have changed, and the staff has been extended. They now mainly work on keeping the databases up to date and on the development of new information systems for libraries and library users. Examples include *Excerpta Informatica*, a periodical plus database system that reviews the literature in the field of information systems and computer science, and the KWIK system, which is based on Carnegie Mellon University's MERCURY system (KWIK is the Dutch translation of mercury), but has been adapted for PC-Windows. KWIK has been developed by Tilburg University in cooperation with Elsevier Science. A WWW version of the system has been introduced recently to make the system

independent of the user's operating system. The system is now based on a client-server approach using Netscape. The library system can be accessed from a local or a remote computer by anyone who is listed in the university's user identification.

A library session starts with entering the university's web site. There the user can choose one of the many databases by using the hyperlinks in the library's hypertext. Here we concentrate on the most important links. First there is a database called Catalogus, which contains information on 852,190 book titles. When entering this database the user also has the opportunity to access the databases of other Dutch universities and the national library. The most interesting databases for this paper, however, are "Online Contents" and "Attent." Attent is a database that contains information on the title, author, keywords, and, in many cases, the abstract of all discussion-paper series to which the University has subscribed. The full text of the discussion papers is not electronically available. We (and many others), however, make electronic versions of discussion papers available for others through the WWW. Currently Attent is a valuable tool to find recent discussion papers related to one's work. To be kept informed a user can ask the library staff to make a user profile. The system then mails to the user on a monthly basis what is new according to his/her profile.

Will such locally developed databases containing information on a limited number of discussion papers on various subjects remain valuable? Specialized databases covering specific areas are also available through the WWW. These databases also contain information on how to obtain the full hypertext, which in many cases can be accessed, reviewed, and, if required, downloaded and printed through the WWW. Examples of such databases are the previously mentioned ERN and CORRYFEE.

THE ONLINE CONTENTS DATABASE

Online Contents is, for a scientist, the most interesting database in our library system. It contains information on title, author, journal, and, if available, keywords and abstract, of all articles that appeared since 1991 in approximately 1,700 scientific journals (out of the university's 2,300 subscriptions). The main criterion for inclusion of a scientific journal in Online Contents is that its articles can be linked to an author (or authors). Furthermore, Online Contents contains pointers that link the article descriptions of almost 120 journals published by Elsevier Science (to which the library has a subscription) to their full text images. The text images themselves are located on a document server. The set of journals of which the full text images are available electronically is currently being expanded with journals from other publishers (Academic Press, Kluwer). Online Contents is remarkably easy to use. The user can search the database using a string of text once he/she enters the WWW. The system then lists all articles that contain this string on the screen. Next the user can select an article and look at the description. If the full text image is available for an article, the user can inspect it on the screen and print it on the local printer if necessary. It is obvious that this is an important service for an academic library. There are, however, a number of drawbacks, which are discussed next.

Not All Articles are Available Full Text

The full-text images of articles published by special interest groups, such as the Econometric Society and the Institute for Operations Research and the Management Sciences, and those of a number of small publishers are not available. From a user's point of view, this shortcoming in the database must be overcome as soon as possible. Small publishers and special interest groups are

reluctant to invest, because it is unclear whether such an investment will pay off. For competition reasons, however, we expect that the full-text images of most scientific journals will become available within the next 3 to 5 years. This, of course, requires cooperation between all those who publish scientific journals.

Availability Affects Your Reading

The availability of full-text images in a database that can be searched, plus the fact that the articles can be printed on a local printer, affects the way articles are used. The most important journals in one's research area are consulted anyway. But people are lazy. The fact that the full-text images of a substantial set of international journals are accessible locally, whereas others require a trip to the central library building plus manual copying, affects your reading. The information on articles whose full-text image is not available tends to be forgotten once the printed full text images of other articles are on your desk. Because the importance of a scientific journal is partly based on the number of references made to the journal, getting into a full-text database will become a competitive advantage once full-text databases are widely used. Making a journal electronically available is now an option, but will become a must to survive in the future.

Quality of Printed Images

When a laser printer is used, the quality of the printed images is good (better than a photocopy). If, however, an article contains one or more figures with color or different intensities of grey, the printed image of such a figure becomes a black hole; no information escapes to reach the reader. The reason for this is that currently the image of an article is a scan of the printed copy, which becomes available several weeks *after* the printed journal appears. To an outsider the current scanning practice seems a bit awkward. It is at odds with the view that IT should improve and speed up the dissemination process. In the near future, however, the electronic version in the form of a PDF file will be at the beginning of the pipeline, not at the end. Much attention is paid to the quality of the printed image. For most users this is, however, of minor importance. If discussion-paper quality can be achieved they are satisfied; it is the content researchers are interested in. They hope that working with electronic versions will reduce the time between the submission of an article and its appearance.

Quality of a Search

Assume scientist A searches Online Contents using the term "energy." This leads to 844 responses. To narrow this number down the set is reduced by using a second term, "prices." This reduces the set to 15 articles, denoted as S_A . Scientist B first uses "energy economics" (258 responses) and then "prices," which results in a set of only nine responses, denoted as S_B . Scientist C immediately searches with the term "energy prices," which results in 36 responses (S_C). To the user it is not clear what causes these differences in set sizes. One would expect the three sets to be the same. The least one would expect is the set S_B to be included in S_A , and S_A again in S_C . The first inclusion holds, but not the second. The intersection of S_A and S_C contains only two titles.

A possible explanation for this is that many article descriptions lack adequate keywords. Currently there is no list of agreed upon keywords used by all who write articles; most authors make up their own keywords. Databases, such as Online Contents, have to be constructed very carefully. As the above example shows, even for a simple concept such as "energy prices," three search routes for which the user would expect the same outcome lead to very different results.

Constructing a database such as Online Contents requires well-defined concepts and discipline from both those who create the database and those who write the articles incorporated. The latter need to be very concise in naming keywords, titles, and JEL codes, to facilitate the work of the former.

Redundancy

A well-formulated database shows no redundancies, that is, does not contain the same information two or more times as a separate item. Redundancies can be avoided. The three sets S_A , S_B , and S_C , however, each contains redundancies, that is, at least one article is contained twice in the set.

NEW OPPORTUNITIES

The WWW already offers many opportunities for scholars and scientists to communicate the results of their research. Electronic exchange of scientific output through centralized and/or decentralized databases is growing rapidly. Personally I gather a lot of information this way. Many scientists already offer hypertext versions of their results, with links to references, data used, and computer programs that can be run or downloaded by other researchers for their own use.

For several reasons the ERN and other databases can become a threat to the traditional form of publishing. First, some groups in academia think electronic publishing is a route to avoid professional publishers whose main interest is to make a profit. They expect that electronic publishing will reduce the cost dramatically [1,7]. An argument in favor of working without commercial publishers is that much of the work (editing, refereeing) is already done (and mostly for free) by scientists, not the publisher. By using the Internet, publishers can be avoided. Second, the time between submitting a paper and when it is published is several years in many research areas. For scientific communication and discussion this is far too long. Especially in the core of one's research, faster communication is required. Normal practice is to write a discussion paper, present this at an international conference, and update it based on the comments received. Then the updated paper is submitted. Electronic communication speeds up this process, and many subject lists, such as the previously mentioned SAMO list, were established to facilitate this communication. Third, academic discussion can become normal again. (The first electronic conferences with contributions from all over the world and using the WWW have already been organized by the World Bank.) It is easy to link hypertext comments to an electronic version of a paper. Of course, much of the time delay in publishing is due to the refereeing process and limitations in space to schedule publications. Electronic publishing does not speed up the refereeing process, but it does not have space limitations. An article can be added to a journal's database as soon as it is accepted. With respect to the refereeing process, changes are possible too. For example, instead of strictly reviewing an article before publication, one can think of less rigorous reviewing, but link a peer discussion of an accepted paper to the text. Publishing articles in the form of hypertext, with links to all important references, data, computer programs, ongoing research, etc., would be considered by many scientists to be a real improvement.

If the electronic research exchange is so marvelous, why do scholars and scientists still want to publish articles? The answer to this question has several aspects. Comprehensive databases do not exist in all areas of research, nor do all scientists already have access to the WWW. But even if all scientists in a field of research had access to the WWW and a comprehensive database, they would still submit papers to journals. In the field of high energy physics, the number and volume of

journals has not diminished. There are several reasons for this. First, an article in a scientific journal has more status. By publishing it as a journal article, a contribution is recognized as a valuable contribution. Second, the status of the journal is a measure of the quality of the contribution. Originality is measured by the number of references to the article, which is more likely to happen if published in the best journal possible for the subject. The publications are needed for recognition and funding of one's research. At Tilburg University (and many others) there is a direct link between the number and quality of your publications and the amount of time and money you can spend on research.

Only electronic journals with refereed papers will be recognized as a respected channel for scientific output. Since it takes at least 5 to 10 years for a new journal to build up some reputation, it will take a long time before the new publishing routes become a threat to the current practice. On the other hand, the fact that well-known scientists from reputable universities are active in building and maintaining databases in different fields of research is constantly improving their status. For the time being, the chance that traditional scientific journals will disappear can be neglected. However, the fact that journals will shortly become available electronically already opens a variety of new opportunities to distribute their contents.

Not much is known about how journals are actually used. To obtain some information on usage for my own faculty, I compiled a list of journals that were in the reference lists of the discussion papers published by the university's Centre for Economic Research (CentER) during the 5-year period 1991–1995. Seventy to eighty percent of the faculty's scientific output in the field of Economics and Management Science/Operations Research is published in this series (80% of these discussion papers were later published in scientific journals). There were references to 377 journals, but to only 82 journals (22%) are there 10 or more references during the 5-year period. To most journals there are three or less references. Of course, the number of references to a journal may not be the best indicator of its use, but if we accept this indicator it is clear that the availability of electronic versions of journals can drastically change the market.

Subscription prices are going up much faster than the average rate of inflation. These price increases in combination with decreasing budgets are forcing many to cut subscriptions. In cases where electronic versions of journals are available, other possibilities than either a full subscription or no subscription at all become possible. For example, the faculty might want a full subscription to the 22% that constitute the most heavily used journals, and pay for other articles by printout and/or usage time only. There is a potential win-win situation. Although publishers will lose money on subscriptions, there is no longer a need to limit the number of journals that can be accessed; every university or institute can use whatever the publisher has to offer. The journals will be more widely used.

REDISTRIBUTING THE RENT

The goal of commercial publishers is to make money; special interest groups that publish journals at least want to recover their cost. There is much speculation about the cost reductions that can be achieved; they vary from 10% to 20% to reductions in the range of 70% to 80%. How much it will actually be is unknown. Who is going to profit from the cost decreases, the users or the publishers? The answer to this question depends on how the publishing business will be redesigned due to the new technologies. The Internet can be used by many, but investing in large databases or electronic warehouses with many opportunities requires much more than access to the Internet. The Internet is a highway, and it is getting congested. (A second Internet for academic purposes only is under consideration.) Future opportunities will depend strongly on how the Internet and the

related technologies develop [8]. The urge to break free from publishers as promoted by the PostGutenberg Galaxy is not based on research, but only on the feeling that, in this way, the accessibility of scientific work is improved and the cost can be reduced. However, what is electronically available now is often directly or indirectly subsidized. More research on this issue is required. Taking into account all costs and benefits may show that the use of professional publishers lowers the cost, despite their profits [9].

But even if there are no gains in the form of cost reduction, there are still plenty of opportunities to get more for your money. Currently our faculty spends almost two thirds of its budget on subscriptions, many of which are hardly used but cannot be missed for reasons of collection completeness. The availability of electronic journals on a remote server, however, reduces the need to maintain a complete collection; the journals are available anyway. This makes it possible to formulate contracts between publishers and libraries that serve the demand profile of each specific library. Because the library system keeps track of demand, a demand profile can, for the first time, be based on the actual use that is made of journals. The number of possibilities are almost endless and can be tailored to the needs of individual institutes. (See also the contribution in this issue by MacKie-Mason.)

A modern library system also offers the opportunity to promote rational use. In the current situation there is no link between subscription cost and usage. Users perceive the availability of a journal as free of charge. As a result, they want a subscription to as many journals as possible. The same goes for printing an image; if this is free of charge they will print too many. Within our faculty, the philosophy is to educate people about the financial consequences of their actions. Therefore, a request to print an article is followed by the question: are you sure you want to print this? In the future this will be supplemented by stating the actual cost of the print, which will affect the user's behavior. Cost-effective usage can be induced further by introducing individual or departmental budgets that are checked by the system. If the budget limit is met, action is required to lift it. Our experience is that once staff members are made aware of the cost of their requests, they reduce the number of requests. Currently there is no agreed upon tariff structure for using and/or printing an image. As an economist I know that there are a number of ways to construct such a tariff. The problem is similar to billing electricity or phone calls [8,9].

The electronic revolution affects the traditional role of the academic library as an institution that collects and stores information and makes it available to its users. An electronic collection is not necessarily present at the library; it can be stored anywhere. Furthermore, for the user there is no difference between what is made available through the library and what is made available through the WWW by others (such as the Economics Research Network). As a result, a new role for the academic library emerges: providing intelligent assistance to its users to cope with the ever increasing stream of information.

CONCLUSION

The Internet in combination with new technologies to electronically store and retrieve large quantities of documents is revolutionizing the way scholars and scientists communicate research results and the role publishers of scientific journals and academic libraries play in the support and dissemination of research results. More and more researchers use the WWW to make their scientific output available to others at an early stage. To facilitate this process, noncommercial databases in combination with mailing lists are constructed by scientists for almost every field of research. Furthermore, the WWW allows researchers to access specialized information resources at institutes all over the world. The books and journals in your local library are no longer the only

source of information. To control the wide variety of information resources, users urgently need assistance in the form of intelligent software systems. The traditional role of the library as a place to collect and store books and journals physically is vanishing with the availability of remote databases containing electronic versions.

The WWW will not diminish the role scientific journals play in academia. The availability of electronic versions of journals will, however, change the way in which they can be made available to libraries. Currently a library either has a subscription to a journal or has no access at all. With the introduction of electronic versions, other ways to sell journals are possible. Users will always want the most used journals to be available at the library. For lower use journals, rapid access plus the opportunity to print an article is sufficient. For the first time in the history of academic libraries it will become possible to order only what is needed and to pay for what is used. Electronic journals in the form of hypertext, giving access to references, data, and computer programs, becomes feasible. Peer discussions can also be linked to an article, which would improve the exchange of scientific ideas. Furthermore, the long delay between submission of an article and its appearance can be shortened considerably. For scholars and scientists these would be valuable improvements. It is unclear to what extent the new technologies will reduce costs. In a time of increasing demand and decreasing library budgets, cost reductions are urgently needed. But even if there are no cost reductions, electronic publishing opens new ways to get more service for the same amount of money. Which new products will emerge, however, is still unclear. This will depend on the market strength of publishers, libraries, and scientists.

REFERENCES

1. Harnad, Stevan. "The PostGutenberg Galaxy: How To Get There From Here." <http://cogsci.ecs.soton.ac.uk/~harnad/THES/>, 1995.
2. Furner-Hines, Jonathan and Peter Wilett. "The Use of the World-Wide Web in UK Academic Libraries," *Aslib Proceedings*, 47 (January 1995), 23-32.
3. Pagell, Ruth. "Planning Your Technology Mix: Decision Factors That Meet User Needs," *The Electronic Library*, 9 (1991), 319-324.
4. Los Alamos, E-Archives, <http://xxx.lanl.gov>.
5. CORRYFEE, <http://tomade.ere.umontreal.ca/mili/ctrdoc/jepl.html>.
6. The Economics Research Network, <http://www.SSRN.Com/>.
7. MacKie-Mason, Jeff and Hal R. Varian, <http://www-personal.umich.edu/jmm/>.
8. "Why The Net Should Grow Up," *The Economist*, (October 19th-25th 1996).
9. MacKie-Mason, Jeff and Hal R. Varian. "Some Facts About Usage-based Pricing," *Computer-Networks and ISDN Systems*, 28 (1995), 257-265.