Surfaces

OH LORD, WON'T YOU BUY ME A MERCEDES BENZ OR, THERE *IS* A THERE THERE

Ann Okerson

Volume 4, 1994

SUR LA PUBLICATION ÉLECTRONIQUE
ON ELECTRONIC PUBLICATION

URI: https://id.erudit.org/iderudit/1064955ar
DOI: https://doi.org/10.7202/1064955ar

See table of contents

Publisher(s)

Les Presses de l'Université de Montréal

ISSN

1188-2492 (print)
1200-5320 (digital)

Explore this journal

Cite this article

Okerson, A. (1994). OH LORD, WON'T YOU BUY ME A MERCEDES BENZ OR, THERE *IS* A THERE THERE. *Surfaces*, *4*. https://doi.org/10.7202/1064955ar

Article abstract

An overview of the current situation of scholarly research and publication on the Internet, with suggestions for further developments.

This article is disseminated and preserved by Érudit.

Érudit is a non-profit inter-university consortium of the Université de Montréal, Université Laval, and the Université du Québec à Montréal. Its mission is to promote and disseminate research.

https://www.erudit.org/en/

OH LORD, WON'T YOU BUY ME A

MERCEDES BENZ

OR, THERE *IS* A THERE THERE

Ann Okerson

ABSTRACT

An overview of the current situation of scholarly research and publication on the Internet, with suggestions for further developments.

RÉSUMÉ

L'auteure offre un survol de l'état actuel de la recherche et de la publication sur l'Internet ainsi qu'une perspective sur les développements à venir.

The purpose of this session is to address the content of information that is being or will be distributed over the Internet, and to consider questions such as "Who is writing in that medium?" and "What are they writing?" Related topics are the impact of preprints for "formal publishing," the question of who maintains bibliographic control and whether there is a Master Index of electronic journals. I also ask why is peer review important or if it is at all.

I would like to develop two points for your consideration. The first is that the content that researchers and scholars are currently placing on the Internet is indicative of the kinds of communications and publications *they* want; and the second is that librarians and some of the learned/professional societies are starting to organize that information in supportive ways that will make it more and more easily accessible.

What Internet content can teach us:

The Internet (like its precursors from the late 1960s onward) offers scholars a unique opportunity to communicate and publish using a revolutionary new technology. In part, it is unique because research and academe had the first go at networked connections, long before entertainment and infotainment could benefit from it. Via this subsidized service, they have had means and opportunity to distribute inexpensively and widely at a time when traditional print publications were increasingly hard to afford and new journals harder and harder to start up. If you are willing to believe that the practically unconstrained way in which academic authors utilize the new academic publishing capabilities might instruct us about what they want and need, then you can imagine the publications currently on the electronic networks as a large goldfish bowl for our study, contemplation, and enlightenment. I want to spend a few moments sketching what are the serious thing on those networks because I believe librarians must take their cue from network users and not from print-on-paper publishers and librarians.

We are in the early days of this new, far-reaching, powerful communications mode. It will take time, experiments, false starts, grand successes, and whopping disasters before we know where the path we are on may lead us. While we can try to shape and direct progress, while we can work together to deploy the technology successfully, its evolution will take time. We have little way to know what developments will come next -- either incrementally, as with gopher and MOSAIC, or monumentally, as with the perfect OCR scanner, artificial intelligence, entirely portable quality displays, and so on. Surely this state of affairs is no different from the early days of scholarly print publishing: the early scholarly book was, by any current standards, as poor and unsophisticated an artifact as our electronic creations will be determined to have been in 100 or even 10 years. At present, we cannot tell how publishing and libraries will be changed; how our roles will have been re-shaped; and how our institutions will change as a result. I often wish I could come back for a day every 50 years, to see what indeed has transpired.

What's "out there?"

Let me offer two perspectives on that question, firts an overall view of the Internet world, and second one of a finer granularity, recognizing that anything that is said or printedin an attempt to capture or define this mutable, fleeting world is doomed to be incorrect the second after the evidence is gathered. The large view comes from a very useful OCLC (Online Computer Library Center) document released in February 1993: *Assessing Information on the Internet*, by Martin Dillon, Erik Jul, Mark Burge, and Carol Hickey of the Office of Research.[1] One of the objectives of the authors' project was to investigate the nature of the textual information available on the Internet, and I am not aware of any other study that would provide a more detailed snapshot of our object. According to the authors, at the time of the study (spanning most of 1992):

* More than three million files representing 165 Gbytes of information were available from 1,044 ftp sites.

* The 20 largest sites accounted for 57% of the available files.

* During the period of study, the number of sites increased by 25%, the number of files by 46%, and the amount of storage by 63%. (Presumably OCLC could use those 8 months of growth data to make a projection for the size of activity a year later.)

* Based on all files at 1,044 sites, 43% of files were source & system code; 12% were news (such as archives of Usenet newsgroups); and 10% were text files.

That is, if growth on these sites has continued more or less proportionally, in the absence of any information to the contrary, we can assume that some undetermined percentage (and probably the bulk) of information on the Internet might be in areas that libraries would wish to organize so as to enhance access. Possibly 10% -- still not an inconsiderable amount -- is currently what we could characterize as an area of interest for both librarians and publishers. At a guess, this 10% would be characterized by electronic book type material, i.e. humanistic texts, corpora of assorted textual files such newspapers, etc., and thus very little of what we would consider current scholarly books. The other chunk would be electronic

serials, i.e., publications which are intended, at the time of creation, to continue indefinitely.

Electronic Journals on the Internet

Our own experience in the Association of Research Librarians (ARL) is predominantly in the serials subset of that 10%. (I wish I could say what percentage -- we should probably encourage OCLC to take on this challenge as well.) As you might know, since 1991 the Association has published a print reference work called the *Directory of Electronic Journals, Newsletters, and Academic Discussion Lists*.[2] The first and second editions were based on the work of Michael Strangelove, a graduate student at the the University of Ottawa and Diane Kovacs, a reference librarian at Kent State University. Way back in the earliest 90s, when gophers were mere garden pests, Michael burrowed and dug through the networks ferreting out electronic journals and newsletters; Diane applied her considerable energies to identifying scholarly and academic discussion lists.

By 1993, the scholarly lists section project team had expanded to 10 Kent State professionals and list inclusion had grown from the original 517 in 1991 to 1152 in 1993 -- in part this growth reflects the growth of academic discussion lists on the networks; in part the expanding mandate of the compilers who added UK Mailbase groups in 1993 and will be adding scholarly UseNet lists for the 1994 edition, again projected to expand its size by at least 30%. (Note that the universe of discussion lists is estimated at a current total of between 4,000 - 8,000, and that many of these lists come and go; many are not archived, so snapshots are the best anyone has done; and those have only been made in subsets of the List world.)

As to the electronic journals and newsletters, after locating and identifying 110 titles in 1991, 133 in 1992, and 240 in 1993, Michael Strangelove began a print serial publication called *Internet Business Journal* and as of summer 1993 became only one member of an expanded and institutionalized E-journal/newsletter directory project coordinated out of the ARL, which will be responsible for this part of the 4th edition due Spring 1994. Currently, about a half dozen workers are contributing to a project which is being re-shaped and re-defined. Our universe for the Fourth Edition, and it is a universe we are researching meticulously to understand what is "out there," includes: the 700+ e-journals on the CICnet gopher (by the way, these are all far from real serials, for various reasons); the previous editions of our own Directory, the list of Zines maintained by John Labovitz at netcom.com, and the 40+ new e-journal startups that have been documented on our moderated e-site: NewJour-L@e- math.ams.org. Our listing of current electronic journals and newsletters as of Spring 1994 will be an estimated

400+ titles. These numbers will exclude the small or proprietary experiments available to selective clientele. I can imagine the day, of course, when every serial in Ulrich's has an Internet format.

What else is there by way of serials content on the networks?

Principally , the rest of the field falls in two areas. First of all, preprints are a comparatively tidy activity. That is, you can find the sites and count them. Through the work of a small Preprints Group (1992) which Dave Rodgers (AMS), Jim O'Donnell (Penn) and I initiated, we attempted to document and report the emerging networked Preprint phenomenon, and hope to include a brief listing of the most prominent preprint services available on the net in our next edition. These are predominantly in the areas of high energy physics, mathematics, and philosophy, with some in computer science, astronomy, and philology.

Second are the awareness, indexing, and abstracting services that are coming on stream. Just how we will incorporate into the Directory such "secondary services" as the tables of contents that Springer and Kluwer distribute over the nets; the astronomy abstracts that the AAS community is making available online; and the alerting service that Institute of Physics Publishing and Elsevier are beginning in 1994 for condensed matter materials, remains to be determined. Maybe we will simply record them as the beginnings of an important and natural trend.

Generalizations:

If content on the networks is to instruct the library community's efforts, what generalizations can one make?

* Something powerful and important is happening. It is growing very fast.

* Some of the things on the Internet are not quite like anything we have experienced in the print world: such forms would be the discussions, the apparently expanding role of preprints, and hybrids that we can't yet quite characterize. For example, one hybrid is the scholarly list linked to fileservers with scholarly articles and texts, and another is the Astronomy

resource developed for the World-Wide-Web by Bob Hanish of the Space Science Telescope Institute at Johns Hopkins University. In his compendium Hanish knits together in one hyperlinked resource various kinds of information sets, such as abstracts, datasets, preprints, articles, journals and software, regardless of variations in format. His resource takes one to each of them by a mere click on the highlighted name of the resource site.

* The low-end electronic publications of today will, in a number of cases, move up scale as public domain technology continues to make new and important publishing offerings every year or so. For example, the earliest and most basic, listserver distributed ascii journals, became available through gophers in 1992 and some of their editors are now tagging them with html (hypertext markup language) and making them accessible through the World-Wide-Web. Which makes them look very pretty indeed, and makes them exceptionally easy to use.

* In spite of some rather superior comments about how low in quality these journals are -- they don't look good, they aren't scholarly enough, they don't have the right imprints or possibly any imprints to speak of and their accuracy can't be absolutely guaranteed - -- these journals at which we might be tempted to turn up our noses, clearly have value and are significant. They attract hundreds and thousands of subscribers. By starting more and more of them, their creators show us a success story, and they do it in spades.

Apparently, there are academics, and reputable ones at that, for whom the cost/benefit of the Mercedes Benz -- the smart cover, prestigious logo, beautiful paper, and added-value galore -- is less important than the means of quick and effective conveyance, even if it be merely a rusty old heap that runs. Academic aspirations are, in many cases, being modified by the financial realities of the day. I believe this is leading us to a more differentiated array of publications. I imagine the Internet full of curiously painted VW beetles and vans, an engaging mixture of information vehicles. If this speculation becomes reality, and if our academics and their institutions become aware that the current style of single-minded high-value publishing *can* lead to perishing, then we are headed for some value shifts over time. I recall that in the first year of the Directory project, we approached a couple of prestigious foundations for support, and were turned down on the basis that the networked content was just not good enough to track and describe. That 1991 judgment has not stopped the phenomenal growth of networked content, however awful some think it is, nor, fortunately, has it deterred our librarian navigators on the project.

Taking the Internet seriously

I shall list only a few significant library community efforts that are coming on stream, advancing networked scholarly resources for locating and describing information . And in so doing, I will shortchange the obviously important work of several of the scientific and humanistic societies who are organizing the Internet resources that their researchers and scholars use (examples include the American Mathematical Society, the American Astronomical Society, the American Philological Association, and others.

* The OCLC report cited at the beginning of my talk made important recommendations about how libraries need to describe the new electronic information -- in fact it stated strongly that the community has no other choice but to do this. At a national level, the LC (Library of Congress), OCLC, RLG (Research Librarians Group) and like organizations are addressing exactly the critical issues of description and organization of networked materials.

* The ARL is creating a Task Force that will formulate a strategy for how the 119 member research libraries should best participate in the housing and dissemination of electronic materials, initially of journals and preprints. We hope we will attract representatives from the national library organizations to our conversations and work together to determine what part we should play.

* The ARL's work on tracking and documenting the location of e- journal and newsletter startups will continue and intensify. I hope we will work with organizations such as LC, CONSER (Cooperative Online Serials Group) and others to lassoo these works and their editors into the national and international world of organized information, by seeking their collaboration in telling us about their publications and where they may be found, by helping us to describe them both bibliographically and through standard numbers, and by assuring their long-term longevity.

* The CIC Libraries and CICnet ("Big Ten" universities in the Midwest) have written an important report committing to a continuing interest, support, storage and access for academic e- journals and text files. This is benchmark work by a major group of research institutions that should be followed and supported.

* The ARL Office of Scientific & Academic Publishing plans for 1994 include offering services for academic authors who wish to start Internet

publications. In this I hope we will join forces with other organizations, such as scholarly societies, with whom we share many common goals.

* Libraries are positioning themselves to work as interfaces between scholarly information -- books, journals, databases, or whatever it may be -- and its users. Libraries are saying to publishers, be they individual faculty author/editors or major publishing houses: we are developing the skills and infrastructures to present Internet information in ways that are user-friendly and functional, that are compatible and integrate well with our campus information systems.

As for librarians, it is in listening to our clientele -- readers and authors -- that our most effective and long-lived future lies.

Ann Okerson, Director
Office of Scientific & Academic Publishing
ann@cni.org

Surface Page d'Accueil/Home Page

[1]*Assessing Information on the Internet; Toward Providing Library Services for Computer-Mediated Communication*. Research Report (Dublin, Ohio: OCLC Office of Research, February 1993).

[2]The data in these paragraphs is taken from the *Directory of Electronic Journals, Newsletters, and Academic Discussion lists*, ed. by Ann Okerson (Washington, DC.: Association of Research Libraries Office of Scientific & Academic Publishing, 1993 edition).