Abstract: If we are to learn about the impact of computers and networks on society, it will be necessary to address issues from broad cultural-historical perspectives, such as has been done for print culture by those in the *histoire du livre* tradition. There are paradoxes faced by users of the Internet that have direct implications on their conceptions of the organization of knowledge. Perceptions of knowledge structures may play roles in searching habits or in deciding about the overall appropriateness of a Net search. The Net has been compared to a world brain and is here placed in the context of an early conception of a world brain.

1. Introduction

Human conceptions of knowledge are based on socio-cultural environments and structures of information sources. Some structures are institutional, some personal, others related to information retrieval systems. When confronted with the Internet and its various interfaces (known also as the Net, or through hypertextual or graphical interfaces as the World Wide Web, or simply the Web), users are forced into an almost chaotic realm of complex relationships between and among sources and into an unpredictable array of searching methods. That the entire process of research, if one can call it that, is disconcertingly difficult to evaluate should be clear to anyone who has used the Internet casually or formally. Whether a search has been exhaustive or even efficient remains ultimately unknowable. Closely related to evaluative measures are users' perceptions of the nature of knowledge, in general and on the Internet; for some users, the Net may serve as their conceptual model of knowledge, its nature, form, likelihood of existing, organization, reliability—in short, a kind of personal epistemology. The "coming of the computer" has had radical effects on society's perceptions of knowledge—perhaps not on the nature of knowledge itself.

There are several questions that should drive enquiry in this area and they are the issues that direct this paper. By stepping back from the positions of contributor or user, it would be desirable to examine the Internet from a broad perspective, perhaps on the order of two landmark works in the *histoire du livre* tradition (Febvre and Martin, 1956; Eisenstein, 1979). Second, it is important to view the advances in networked communication and computing as reflections of seemingly paradoxical realities, which may have direct effects on perceptions of the Net and which may help us understand more about the impact of networked computers on society, particularly on conceptions of the nature of knowledge. We are not necessarily dealing with knowledge itself here; rather we are concerned with people's perceptions of it. Third, insofar as we may be concerned about the ability to "know" only that which is "true," we should be concerned with evaluative measures of Internet sources and the Net itself as an information retrieval system, or set of retrieval systems. This paper addresses some issues of information retrieval, particularly the feasibility of measurements of recall and precision. And lastly, it may be useful to place the Web in a context of one particularly perceptive—and early—approach.
toward the conception of a world brain. At the risk of paralipsis, a necessarily superficial treatment of this topic, much of significance has to be omitted.

2. The Coming of the Computer and Networked Communication

To what degrees have computers and networks had an impact on society? In certain arenas the effects have been profound, in others there has been no effect at all. To paraphrase Febvre and Martin (1958, 12): The story is about something other than the history of a technique. It has to do with the effect on late-twentieth-century culture of a new means of communicating ideas within a society that has been essentially aristocratic, a society that has accepted a culture and a tradition of learning which has been restricted to certain social groups. An elite has once again been served.

The first users of information technologies have always been members of an elite of some sort, whether from scribal cultures, the world of early printing in Europe, or in the Web culture of today. Tied closely to literacy in each case, the elite also communicated in non-vernacular tongues. The printing press was an agent of change in this regard: "The unified Latin culture of Europe was finally dissolved by the rise of the vernacular languages which was consolidated by the printing press" (Febvre and Martin, 1958, 332). In the case of computers, networked or not, there has been said to be a *lingua electronica*—or perhaps several—(Compaine, 1983, 18) that was understood by a small group at first, then by more, to be replaced ultimately by vernaculars.

In scribal cultures, oral transmission was highly valued by literate elites (Eisenstein, 1979, 11). Eisenstein maintains that in such an environment, learning "was governed by reliance on the spoken word—producing a hybrid half-oral, half-literate culture that has no precise counterpart today" (Eisenstein, 1979, 11). With the Web, the hybrid may have been realized: great portions of the Internet consist of newsgroups, discussion rooms, bulletin boards, e-mail--mostly unfiltered, unedited, quasi-oral, and some quite illiterate. Other parts represent the cutting edges of scientific and humanistic scholarship--scrupulously edited, carefully selected, and understandable by only the most literate of the relevant disciplines.

The relationships between the Internet and various levels of society would be as useful to investigate as have been the relationships between printing and society. There are of course many specific subtopics worthy of attention. An evaluation of the sources to which the Internet is connected would be in order, but would be much more complicated than categorizing the items listed in Hain’s *Repertorium bibliographicum* or the *Gesamtkatalog der Wiegendrucke* into disciplines: the sheer number is prohibitively large and growing; perhaps thousands defy categorization. We have numerous studies of incunabula that tell us much about printed output before 1500, but what have we for Web sites? Ephemeral "Internet Yellow Pages" may help. To be sure, Hain carried out his research centuries after the period he studied. What will remain of today's Web in 400 years, or in 10?

How has the Net preserved the moral, religious and literary heritage accumulated by its predecessors from the eleventh to through the twentieth century? How has it ensured the continuity of a tradition which has linked us with the past? Conversely, how successful has the Internet been as an agent for the propagation of new thoughts? What role has the Net had in the contagion of political ideas? How has it, by design or accident, helped spread English around the world even more than it already has been? Are there other relationships between language and the Net? What effect has the Net had on education, business, tourism, or personal enrichment?

Cultural bibliographers have studied the geography of the book. An historical geography of the Internet showing changes, perhaps on a yearly basis would be enlightening. An
historical Net demography, describing characteristics of the population of Net users and contributors would likewise be useful.

A reading history of the Internet could be fascinating. What literary forms might it claim responsibility for spreading? How has hypertext changed the way readers or authors think about linear texts? Which literary forms has the Net resisted? Clearly within bounds of cultural bibliography would be studies of Net economics. To what extent does trade exist on the Net? How have conceptions of profit changed since the early years of printing? Some scientific works with limited market appeal existed in manuscript long after printing was introduced. On the Internet, items with very low market value, some would say with no market value whatsoever, can be mounted easily by individuals. How can popularity be measured? For some printed editions it is possible to determine the size of print runs; it can be more difficult with Net sources. How often is a source accessed, read, or copied? Although some Web sites maintain usage statistics, many do not. Who has access to the Internet? Has it had an impact on all of society or only on parts? Printing had an effect on an elite with a subsequent effect on larger groups. Does the Internet reach larger groups in the same way? What are differences in age, gender, academic discipline, socio-economic status, and level of computer literacy that can tell us something about the relationship between the Net and society?

Modern researchers face a major problem not encountered by Febvre and Martin or Eisenstein: the Net has not been around a long time. Is it possible to gain an objective perspective when one is so close to one’s subject? For example, Febvre and Martin examined the relationship between printing and the spread of religious ideas between Catholics and Protestants, and other later groups: Freethinkers, Deists, Atheists, and Materialists. If such relationships exist between the Net and the spread of religious ideas today, they are difficult to observe. We are faced here with a conundrum: we do not have the perspective of time that can be so useful (conversely, we do have the advantage of being eye witnesses), yet it is all the more important to record these things now because of the ever-changing nature of the Net.

There is a certain ambivalence in the Web subculture toward the cultures it believes it has replaced. This follows a pattern from the Akkadians, who replaced the Sumerians, to newly-elected politicians of today, who often want to distance themselves from their predecessors. For this reason alone, it is not too early to call for an updated approach to cultural bibliography, a new incarnation of the histoire du livre movement, in order to come to terms with the impact of this revolutionary force on society.

3. Web Dialectics and Perceptions of Knowledge: Surprises and Disappointments

It would be unwarranted to adopt a metaphysical system that attributes an empirically sound reality to perceptions of objects or of environments, physical or electronic. Knowledge is different from perceived knowledge; the structures of recorded knowledge are not the same as perceptions of such structures. Yet it can still be a useful exercise to wrestle with paradoxes of the Internet that are present in a working conception of the Net, and likely present in an objective reality. The following propositions would seem absurd in most information retrieval systems:

- The Net or its sources are permanent and ephemeral;
- static and dynamic;
- intuitive and counterintuitive;
- quick and slow;
- predictable and unpredictable;
systematic and unsystematic (yet not random); organized and entropic.

These dichotomies exist on the Net, are apparently mutually exclusive, but may in fact express truth. Some of these may influence the construction of users' mental models of the organization of knowledge. The Internet and the electronic sources to which it provides access are not knowledge; they do not know anything. The Net and the sources together amount to a large information retrieval system. Here we are interested in how recorded knowledge or communication have changed since the advent of the Internet and in how perceptions of recorded knowledge or communication have changed. We may be puzzled about this set of epistemological problems: the nature, scope, limits, and perceptions of human knowledge in an electronically networked society. There are at least two assumptions made in this paper about perceptions of Net knowledge. First, in using the Net we are limited to almost no "knowledge by acquaintance" except for direct knowledge of the Net and its interfaces. Second, while it is true that humans are limited to "true" knowledge because one cannot have "knowledge" of that which is not true, perceptions of knowledge, such as those that Net users may have, exist independently from the presence of truth. People hold beliefs for a variety of reasons, of which empirical thought is only one.

At early stages of some users' acquaintance with the Net, they may not have formulated a formal mental model of knowledge organization in that system. At this point they might describe the Net and its resources as disorienting, overwhelming, confusing, or amazing. In time, users may well continue to use these descriptors, but they will have acquired a more structured model of knowledge as represented on the Net.

Does such an orientation, rather than one of common conceptions of knowledge based on academic disciplines or on divisions within a company, have an effect on individuals, groups of individuals, or perhaps even a generation? Two people using the Web have different perceptions, which are due to differences in the structuring of what is experienced. How does this structuring take place? Any one person uses a small subset of the Net, and acquires over time a personal perception of its organization and contents. As most of us have acquired a knowledge of linear texts, some users of the Net have become comfortable with non-linear texts, with reactive or interactive texts. These can certainly be experienced in different ways. They are so designed. Some users may feel the Web is a random arrangement of items and acquire a particular sense of knowledge organization based on that perception. Like them, Gulliver ran across an encyclopedic frame with random words "in all moods, tenses, and declensions," which would ultimately produce "a complete body of all arts and sciences" and would work even faster "if the public would raise a fund for making and employing five hundred such frames (Swift, 1735; pt. III/V).

A "web" is an apt description of the network portion of the Internet. It is a useful conceptual model that reflects its hypertextuality; however, the analogy is somewhat inaccurate. The Web is not confined by a perimeter like a typical spider's web. Rather than being two dimensional, as many spiders' webs are, the Web is multi-dimensional. The Web is less static than that of a spider; it changes considerably over time. Web sources are not at all well-served by this model: at a given site, one may find practically nothing (perhaps a reference from an old to a new URL) or one may find the electronic resources of the Library of Congress.

One may wonder if the Net is randomly ordered or highly organized, but seemingly infinite. Indeed, conceptions of knowledge such as the random word frames observed by Gulliver or the infinite library of Borghes's "universe (which others call the Library)" (Borghini, 1956) can interpret knowledge as something with much, or without any, order. A similar
situation exists with some music from the middle of our century: that which is highly serialized (music in which durations, attacks, dynamics, and pitches are all rigorously organized) and that which is highly aleatoric (music in which many performance decisions are left to chance).

4. Evaluation

Perceptions of Net knowledge structures may play roles in searching habits or in deciding whether a Net search would be appropriate. How does one decide which way to search for information? In order to satisfy my curiosity about the ability of the Web to provide information on an obscure topic, I called upon one of the many search engines (Lycos in this case) to seek documents about Adalbert Blumenschein, an eighteenth-century Austrian librarian and priest who wrote a stupendous, but little-known, guide book to libraries of central Europe. Expecting to find nothing, I was pleasantly surprised to find two items: an abstract of an article in a scholarly journal at the journal’s Web site and an abstract of a paper delivered at a conference in Edinburgh at the Web site of the learned society sponsoring the conference. I don’t know of a library system that would be as successful. The results were also disturbing; it was only a happy coincidence that the journal and the scholarly society had Web sites and had this information available in a searchable format. Think of the thousands of publishers and other organizations that have no Web presence. It is not systematic in this regard.

The search engines (such as Lycos and many others) and classified arrangements (such as Yahoo and others) are the most important means of searching the Net. Both types of access were designed and made accessible many years after the Internet had existed. The Internet was not originally designed to be an information retrieval system. For this reason, it is not surprising that some established means of evaluation do not work well.

One measure of a system involves identifying and evaluating its sources of information. A survey of Web documents may tell us much about the system’s vastness. Evaluating the sources to which the Web leads is a monumental task perhaps best accomplished by focusing on specialized topics. Will either of these tell us how sources are actually used? No. Plays of Aeschelus are easy to find and it is always nice to know they are there; however, how many people actually read them in that form? The evaluation of individual Net sources is an overwhelming responsibility that has been aided by the presence of lists of recommended sites selected by librarians and other experts, as well as by Net source cataloging programs.

Evaluations of the Net, if they are applied at all, should be made not only to the search engines, but also to the classified systems, including some virtual reference libraries, the Yahoo index, and some self-described encyclopedias. I never felt the loathing of the Web’s Yahoo Index that Gulliver felt toward the Yahoos he met; being a former indexer, I succumbed immediately to the attraction of the Yahoo index, and would agree with Gulliver that I cannot deny that I am "a real Yahoo in every limb and feature" (Swift, 1735, pt. IV/I).

Are measurements of recall and precision feasible here? Recall, a ratio used to describe the ability of a system to retrieve a percent of relevant documents from all relevant documents in a system, is not useful here because the total number of relevant documents is not knowable. Documents are very fluid and changing. Being part of such an unsystematic system, the presence and quality many sources depend on the sustained interest of an individual or organization. For instance, in a reference class, I had students use the Web to locate reproductions of "The Scream" (or "The Cry") by Edvard Munch. One particular site was especially fruitful: a person, obviously fascinated by the different forms of this image, had created a "Scream" site, which provided links to many digital versions of the work. This is the work of a volunteer who may be excited about Munch this moment, but who may not even
have a Net account tomorrow or in ten years. There are thousands of such passionate collectors. Although students found some images easily, it is impossible to know how many they missed.

Precision, which describes the ability of a system not to retrieve irrelevant documents, may be a more satisfying measure. Many users have already waded through a considerable amount of Web "trash" and have already carried out informal measurements of precision. Another complication, commonly encountered in recall/precision measurements, is that it is difficult to define relevance because it has always been very personal. Measures of recall and precision depend on relevance, which can be so unpredictable, so subjective, that it is difficult or impossible to verify. For one individual, a document or information source may be "close enough" to a subject or "good enough" for a particular use, even though it is not the best or even close to the best. If a fee is required, as they are for some of the highest quality Net resources, will a given user be less likely to use it? If it is necessary to register, will a user be less likely to use a very good source, even if no fee is involved? Will a frustrated user switch topics rather than carry out an exhaustive Internet search? Will a novice user take the time to learn how the different search engines work? Several Web search engines allow for Boolean searching, but in at least one case, the default operator is "OR," which of course has the potential of delivering results that are hardly precise. These problems are addressed in part by variations of the Principle of Least Effort: a solution will be judged by a user to be satisfactory if it is easily found, even if it is not the best available solution, and perhaps even if it is not a solution at all.

5. Conclusion: The Web as World Brain

The "hook" of a recent article in a popular information technology magazine reads, "Given that the Web itself is becoming the sum of the world's knowledge, isn't putting the Encyclopaedia Britannica online a spectacularly useless thing to do?" (Rossney, 1995). The author demonstrates that an online Britannica is indeed an especially useful resource, but treats as axiomatic the "given" that the Internet is assuming the role of world brain, a useful model that has been proposed from time to time.

Almost sixty years ago, H.G. Wells proposed the establishment of a world brain (Wells, 1938). He could not have predicted the existence of cyberspace or hypertext documents or e-journals; he would not have been able to imagine search engines of various capabilities. Yet his proposal seems in some ways to have been realized. Wells's proposal and the Internet go beyond the centuries-old idea of universal bibliography. Bibliographers of the past, such as Konrad Gesner, would surely be astounded by the breadth and depth of the Internet; yet some of them would be indignant over the haphazard way it has developed and about the low editorial standards of many of the sources it provides. What might take the "universal bibliographers" a time to realize is that the success and rapid evolution of the Internet are due in great part to the relatively disorganized way it has been formed and to a practice that has allowed any of millions of users themselves to contribute texts, graphics, other datafiles, collections of materials, and correspondence and to gain intellectual access to them in powerful ways. Increasing attention is being directed toward quality control, and equally importantly, to organized intellectual access.

In his discussions of a "world brain," Wells expresses a dream that is currently being realized. His words may be considered verbatim (1938, 20-21):
This World Encyclopaedia would be the mental background of every intelligent man in the world. It would be alive and growing and changing continually under revision, extension and replacement from the original thinkers in the world everywhere. Every university and research institution should be feeding it. Every fresh mind should be brought into contact with its standing editorial organization. . . . It would do just what our scattered and disoriented intellectual organizations of today fall short of doing. It would hold the world together mentally.

Wells saw the world brain as a tool in scholarly communication, presently one of the most heralded functions of the Internet (1938, 24):

To [the specialist] even more than to the common intelligent man World Encyclopaedia is going to be of value because it is going to afford him an intelligible statement of what is being done by workers parallel with himself. And further it will be giving him the general statement of his own subject that is being made to the world at large. He can watch that closely. . . . He will be able to criticize the presentation of his subject, to suggest amendments and re-statements.

Likewise, and in response to Wells's proposal, Smith agreed that "in a way characteristic of the spirit of the whole, the Living Encyclopaedia would turn the intellectual organization of whatever unit of society it had come to serve into an organic community activity rich and joyous with the spirit of mutuality" (1941, 60).

As concerned about the organization and structure of the world brain as we are about the topography of the Internet, Wells suggested that an Encyclopaedia Society be formed to survey the available material, which he considered to be in "a state of impotent diffusion" and to assemble authoritative subject bibliographies, perhaps even a master bibliography, and to form a general editorial board and departmental boards (1938, 27-28). At the same time, he proposed that the project represent the entire world (1938, 74):

So that while I believe that ultimately the knowledge systems of the world must be concentrated in this world brain, this permanent central Encyclopaedic organization with a local habitat and a world-wide range . . . nevertheless I suggest that to begin with, the evocation of this World Encyclopaedia may begin at divergent points and will be all the better for beginning at divergent points.

It is an understatement to say that "divergent points" characterize the Internet. Smith elaborated by suggesting that in order to avoid the destruction of intellectual freedom in such a vast and diverse plan, "the users themselves must in the ultimate be the controllers" in a situation in which contributors would heed the users to a greater degree than they would heed their readers in the traditional publishing world (1941, 61).

Very well aware of the developments of documentation projects in the 1930s, Wells emphasized the availability of a variety of materials for a wide readership (1938, 76-77):

It seems possible that in the near future, we shall have microscopic libraries of record, in which a photograph of every important book and document in the world will be stowed away and made easily available for the inspection of the student. . . . The time is close at hand when any student, in any part of the world, will be able to sit with his
projector in his own study at his or her convenience to examine any book, any document, in an exact replica.

Wells's proposal has of course not been completely realized. He believed that the new encyclopedism "should consist of selections, extracts, quotations, very carefully assembled with the approval of outstanding authorities in each subject, carefully collated and edited and critically presented" (1938, 20). While there is certainly a large and growing mass of texts and other sources, they are not always carefully assembled, uneven or no attention is given to textual quality, and subjects are not systematically represented. Further, Wells suggests that the system "would not be a miscellany, but a concentration, a clarification and a synthesis" (1938, 20). As it now stands, the Internet—arguably for the better—is a monumental miscellany, is regularly diluted, is by no means clear, and synthesizes nothing.

The Web may resemble human thought more completely than Wells could have envisioned. Like humans, the Web possesses long-ternn and short-ternn memory. On it can be found the unorganized rantings of hysterical individuals and the well-selected and sublimely written and organized texts of some of the best minds the world has seen. Also represented is the dark side of the human psyche. After carrying out a search or two—for purely clinical purposes, of course—I can conclude that there are a considerable number of links that guide one to the nether regions of human depravity. I was not completely ignorant about certain aspects of human behavior, but now I feel considerably less naive. The Net is indeed an educational tool.

Many people consider the Internet to be a living, growing, world brain-like organism with a life of its own. Upon reflection, it may well be living and growing, but its life, which is not self-sustaining, depends on many factors. We know its health will be dependent on the existence of conscientious hosts and users; we know less about the future of its political, economic, educational, social, and cultural life. The need for research related to electronic networking and networked resources is manifold. It is important to develop storage, retrieval, and communications technologies. It is as vital that we understand the organization of the networks themselves as it is to come to terms with the range of sources present on them. There exits a social imperative for us to manage the Internet and its successors within the contexts of its economic and political environments. Likewise, we are obliged to understand the related issues of accessibility to networks. Many of these needs fall into the categories of applied and theoretical research. Several journals, including Internet Research: Electronic Networking Applications and Policy, reflect the need for serious scholarly attention. Especially needed are perceptive examinations from historical and philosophical perspectives.

References

