Scholarly Collections on the Web

Media Reconfigurations at Play[1]

Rune Dalgaard

With the Internet a new medium has become available for communicating scholarly texts. This article focuses on the World Wide Web (the Web) as a global, hypertextual archive for scholarly texts and its significance in reconfiguring the corpus of scholarly texts. The first part addresses the Web in light of hypertext theory and a media theoretic perspective, concentrating on its roots in issues related to the flood of information and its qualities as a medium. The second part will zoom in on the actual use of the Web as a publication and archival medium for scholars, with a focus on two different scholarly archives. The article concludes with some general reflections on the Web as an archive of archives based on the concepts of network and complexity.

A Media Theoretic Perspective

What does it mean for our way of organizing and navigating in the scholarly corpus of texts when scholarly texts move from the printed medium to a digital, hypertextual, global network? In the following I will try to shed light on some of the media-theoretical and knowledge-related aspects of the Internet as a growing global archive for communicating scholarly texts. If the Internet is bringing about changes in our system of knowledge, it is not the first time in history that a new medium has had such an influence. Viewed in terms of media history, one can draw parallels to the significance of writing when introduced in oral cultures where oral transmission of knowledge had previously been the principal form of communication (Ong 1988). Likewise, the printing press is

Dalgaard, Rune. "Scholarly Collections on the Web: Media Reconfigurations at Play."
HUMAN IT 7.2(2004): 138–170

often described as a revolution in the cultural history of the West (McLuhan 1962). Mechanical reproduction increased the access to texts and ensured their transmission by spreading many identical copies of the same text (Eisenstein 1979). Such approaches and interpretations of the significance of media throughout history are now as relevant as ever before, since we are experiencing the emergence and spread of a new culturally dominant medium, the computer.

There has been a tendency to frame questions of new and old media in terms of replacement and radical opposition, as if the birth of one medium is necessarily the death of another and the use of a new medium must be characterized by a radical break with the uses of earlier media. There is reason, nevertheless, not simply to oppose new and old media as though a new medium necessarily replaces older media with something completely different. If the book as a cultural form had a voice of its own it would easily be in a position to pronounce "the rumors of my death have been greatly exaggerated" even though its disappearance has been foretold for decades now. In fact, media history shows that new media seldom merely replace existing media, but more often overtake certain functions in ways that are much more complex than mere replacements, and in this process both imitate and transform existing forms of communication over longer periods of time.

Therefore, this study will not repeat the discussion about the death of the book, but rather focus on some of the similarities and differences we can observe between the typographic system of knowledge and the growing system of knowledge on the Internet, with a focus on scholarly documents and collections on the Web. As we shall see, these relations do not have the character of continuity, nor of a radical break, but are rather a kind of changes that are better understood as complex reconfigurations of material and symbolic relations and remediations of existing media forms (Bolter & Grusin 1998).[2] One could be forgiven for expecting the Internet to provide a seamless and convergent system as this is often implied in the promises of seamless navigation and access to all the information in the world. Not quite so, as the staggering amount of documents and collections on the Web, even though only a fraction of all information in the world, are characterized by increasing complexity in the organization of the scholarly literature. Although initially concept-

ualized with the issue of information overload in mind, the Web presents its own unavoidable issues of overload and complexity. Human judgement and capability of making selections is needed as much as ever when navigating the Web. However, some intermediate steps are needed to support this conclusion.

From Memex to the World Wide Web

Ideas of digital document archives have been around for a long time. To a great extent, they grew out of dissatisfaction with the systems available for managing scholarly texts. In the first part of the twentieth century a number of people became aware of the "flood of information" as a problem for scholarly communication. This happened concurrently with the growing importance of science and theoretical knowledge for providing technological modernization in various areas, which made information overload, as it is now called, a concern of general nature.[3] The problem was perceived to be that the amount of scholarly texts was growing at a rate that made it difficult for libraries to manage and organize, and just as difficult for researchers to orient themselves in. There was thus a widespread impression that scholarly communication found itself in a *control crisis,* to borrow a concept coined by James Beniger (1986). This feeling of crisis appears to have been closely related to the deep mistrust that the World Wars brought towards the belief in progress and modernization.

Some thinkers envisioned new institutional structures to deal with this issue; witness for example the classic idea of a "world encyclopaedia" or even a "world brain" proposed by H.G. Wells (1967/1936). What Wells had in mind was a collective body responsible for collecting, filtering and synthesizing the flood of specialist knowledge in one huge constantly updated encyclopaedia. Others entertained a hope that the problem could be solved or at least handled technologically by developing a substitute for the technology that has been the foundation up until recently: the printed text and the handling of it in the modern research library. Vannevar Bush belongs in this latter group favoring a technical solution.

When Vannevar Bush published the now canonized "As We May Think" in 1945, he had this control crisis in mind. In the article, Bush

formulates the idea of a "Memex:" a personal mechanical archive based on microfilm technology that should make it possible to automatically call up each text in the archive. Ideas of replacing print with microfilm and of automation in the libraries are not unique to Bush and are brought up by several others in the same period (Birdsall 1994). But Bush is alone in proposing that one should be able to make links from any passage in the archive to any other passage or, in other words, a precursor to the idea of hyperlinks. There is in other words a basic vision of augmenting the ability of scholars to traverse the scholarly literature by highly interactive mechanized means such as the hyperlink, although Bush does not use that term himself.

Bush is widely and rightfully credited for his visions in this field (see for example Nyce & Kahn 1991), but not all of his speculations have the character of fulfilled prophecies. An example is his idea that trails of links between passages in a text corpus could potentially form new kinds of books closer than contemporary forms to the way people make associative leaps of thought. This idea has since been taken up by a number of hypertext theoreticians as ideal for a kind of associative network that is claimed to constitute a break with the book as a form of communication (Nelson 1987/1981; Bolter 1991; Landow 1992, 1997). This interpretation of hypertext, however, is not without problems in terms of both its basic assumptions and desirability and feasibility of realization. It is possible that we often develop arguments through associative leaps, but it does not follow from this that an approximation of this creative process should constitute a well-suited form of communication. One can also advance the objection that a sequential trail of links through a corpus of text passages does not differ from a raw collection of quotations. We normally expect more of a text, for example that a scholarly text not merely cites other texts, but also that the author places the citations in context and forms an argument. In principle, of course, one cannot generally exclude the possibility that forms of communication will emerge that make greater use of our associative faculties and that the text as a form of communication will be replaced. So far, however, such expectations have not been met and the most powerful legacy left by Bush is the basic idea of the active link as the core of a document archive. Of course, in Bush's vision the active character of the link was supposed

to be realized by pre-digital mechanical technology and the idea of digital document archives did not appear until later.[4]

In terms of exploring and formulating coherent visions for digital archives, James R. Licklider's *Libraries of the Future* (1965) is one of the first explorations of that theme. Licklider considered the book an awkward medium for storing, distributing, accessing and managing technical-scholarly books, and presents an extensive study of the potentials in a digital library. To a greater extent than Bush, Licklider focuses on the possibilities of automating information retrieval and processing – today an area often associated with research in information retrieval. As in the case of the Memex, Licklider's ideas were visions and were never implemented – in fact he seems to have had too high expectations on the analytical capabilities of computers – but his work belongs to the canonical literature of information retrieval. In passing, we might note that Licklider, in the 1960s, was also a central figure in the milieu that developed the ARPAnet that would later develop into the Internet, which in turn is a precondition for a global hypertext system such as the Web.[5]

Noticeable among those who became inspired by Bush's ideas are Douglas Engelbart and Theodor Nelson who, at the beginning of the 1960s, independently of each other, started to think about developing digital hyperlinks. Engelbart's perspective on the computer was influenced by the idea of "augmentation:" that the computer should expand but not replace our interaction with information. Whereas the thought process in AI (Artificial Intelligence) was largely focused on the computer taking over a number of functions from human information management, Engelbart's ideas were more focused on developing the computer as a *tool* and *medium* for the human management of information (Engelbart 1963). Engelbart was behind not only the development of the first implementation of hyperlinks in the computer medium, but also other revolutionizing interface technologies such as the "mouse," the graphic interface and, as part of this, the idea of "windows" (not to be confused with the operating system) (Engelbart 1984).

At the same time, Theodor Nelson was the first to formulate the concept of hypertext, establishing the concept via an opposition of the linear printed text and the non-linear digital hypertext; a dichotomy that

has since, for better and worse, influenced quite a few theories about hypertext (Nelson 1987, 1/29).[6] The first articulation of the idea of a global hypertext system is normally attributed to Nelson and his vision of Xanadu, which was conceived as a hypertextual archive with a built-in, centrally operated "royalty" system (Nelson 1987). Projected as nothing less than a total publishing and archival solution that could manage dynamic texts, different versions of texts, comments and addenda, Xanadu was an incredibly ambitious idea. It was envisioned as a radically integrated system where an advanced link structure made it possible for a text to be stored in only one place in the system. A quotation, for example, would consist of a link to the stored part of the original text, and for the reader appear as an integrated part of the referring text. This and other advanced link-functionalities that were part of Nelson's ideal global hypertext system have never been realized as envisioned, and Xanadu, just like the Memex, today seems destined for the history of ideas rather than for the history of actual technologies. It is, however, an expression of a powerful set of ideas, highly relevant to our understanding of hypertext, digital collections and the Web.

As is now generally known, Tim Berners-Lee in the period from 1989-1991 developed the World Wide Web (the Web) and based the system on the Internet, which in the same period became adopted as the *de facto* standard for network communication (Berners-Lee & Fischetti 1999). The Web, of course, is just one of countless actual hypertext systems, but from a cultural perspective it is the one that matters, because it has become the dominant interface to stored material on the Internet. The Web in other words realized the connection of already existing visions of hypertext with an actual global digital network. The World Wide Web can, in certain respects, be regarded as a hybrid between Bush's idea of a Memex as a personally automated and hypertextualized archive and Nelson's idea of Xanadu as a global, coherent publication and archiving system with a certain degree of central control and assurance of consistency. The World Wide Web employs the global scope of the Internet, but also its distributed structure and the absence of an overall editorial control or organization. In this respect it can, on the whole, be regarded as a meta-archive or, as Jerome McGann has described it, an archive of archives (McGann 2001, 58). In a media

historic context, Niels Ole Finnemann has pointed out that the Internet's role as a global, distributed archive constitutes one of the most significant actualizations of the general potentials of the computer medium (Finnemann 1998).

In light of its roots in the wish to manage the flood of information, the World Wide Web demonstrates a paradoxical complexity in the amount of forms of communication that it offers, and in the intricate connections between these. In fact, it may be argued that the Internet has both intensified the flood of information and increased its visibility. For this reason, a key to understanding the significance of the World Wide Web lies in studying how the amount of information is organized and managed in this new medium. However, first we shall take a brief look at the medium itself.

The Internet from the Viewpoint of Media Theory

With the World Wide Web the Internet gained broad recognition as a publishing medium and archive. There is thus good reason to briefly compare some of the features of the computer and the Internet with the system of knowledge we have built up around the printed medium. As analyzed by Finnemann (1994), the computer's qualities are so general that it can emulate a wide range of previous media (radio, newspaper, TV, library, post system) and represent a wide range of forms of expression (writing, numbers, images, sound). In addition to this, as a medium the computer has an open potential as it is – to use a popular term – programmable; its functionality is to a large extent not defined by hardware but by software. Finnemann places the decisive media-technological innovation in the invention of a new "alphabet," the digital code, which he defines as informational notation (and later as a binary alphabet). The digital code thus marks an innovation in the media-technological management of the relationship among the material, representational and functional qualities of a medium (Dalgaard 1999). Previous media-technological innovations, such as the invention of writing, the printing press with movable type and the analogue electronic media, can likewise be viewed as a new departure in how the relation between the dimensions of materiality, representation and functionality is dealt with. These innovations changed the configuration of time and space as well as the

interactive potentials that previous media had established for our forms of communication. Let us therefore take a closer look at the qualities of the Internet with respect to time and space, and at its interactive potentials for texts and archives.

First of all it should be noted that whereas it is common to talk about the *potentials* of the computer and the *limitations* of other media, there are plenty of reasons also to address some of the limitations of the computer compared to other media. Despite many predictions to the contrary, up until now digital texts have not replaced paper and the book as a *reading medium.* Rather, the amount of texts that we print and publish has grown with the computer. The attempts to make so-called e-books have in many respects merely had the effect of illustrating that the print text is an incredibly flexible medium that we can navigate in quickly and effectively. Bookmarks, page format, lay-out, indices, comments in the margin and a range of material qualities of the codex volume give the reader a range of different ways to jump around in the text, rediscover passages, and maintain a general orientation in the text – a flexibility that has often been left out in comparisons between the digital and the print text. Thus, whereas it is tempting to view the debate over the book and the computer as a discussion between technologists and book-lovers, it is perfectly plausible, and probably more reasonable, to conclude that so far print has simply been the superior technology as a *reading medium*.[7] We have, however, already become accustomed to *producing* texts digitally, and we are increasingly also starting to *store, organize, disseminate* and *navigate* texts digitally. The Internet, and especially the Web, is important particularly in this context as it constitutes a new distributed storage medium that promises a reconfiguration of our text collections and transforms our ways of navigating the scholarly literature.

Time and Space

When we represent a text in a computer, it is not, like in the case of print text, fixed in a delimited physical object. This is not because a digital text is "immaterial" or "virtual," but because it establishes a new relation between what is stored and what is displayed. It is stored as electromagnetic signals according to the logic of the binary alphabet that enable the text to be copied, moved, searched for (or in), and "linked" directly to other

texts without us physically carrying out these operations. In other words, the end of the fixed relation between medium (the computer) and message (the text), paves the way for allowing the medium to execute a series of different kinds of operations on the text. Connecting computers in a global network transfers this potential from the individual computer to the network at large, giving the text and the archive a different relation to time and space than the book and the library have. The print system of knowledge is built around the dissemination of many uniform, physically stable copies of a given text; a system with the twofold role of spreading the text in space and guaranteeing it over time. Let us start by looking at how the Internet reconfigures the text's and the archive's relation to the space dimension.

On the Internet, a text (or images and sounds) in principle only needs to be stored in one place to be readily accessible from any other place on the Internet. Provided that one has access to the Internet and can/will pay, any locally stored collection becomes global and immediately accessible. Access to texts (and archives) is thus no longer connected with their physical location, but is instead connected with their symbolic address on the Internet. Or in other words: the name of the archive and its web address are significant, while its actual geographic location in terms of access is irrelevant. This ending of the connection between place and access has for instance given rise to the appearance of distributed scholarly portals (sometimes called subject gateways) and journals that organize collections of links to texts distributed over the entire Internet. The dissemination and reproduction of texts thus no longer has any particular value attached to it. However, the editorial functions linked to selecting, organizing and presenting information acquire increasing significance because both the amount and the accessibility of information are on the rise.

The dimension of time raises a different set of questions, especially in connection with securing stability over longer periods of time. In contrast to a storage medium like the book, we do not have direct access to digital texts without a program that can decode and present them for us on a screen or some other interface. While digital copying itself occurs without loss, the accessibility of digital archives therefore requires that the texts be continuously transferred to new storage media and new

formats concurrently with the development of software and the physical disintegration of storage media. Furthermore, with respect to organizing content, the Internet raises many new editorial questions regarding the organization of collections. For instance, it no longer suffices to maintain isolated documents and collections, as their relations to other documents and collections through, for instance, hypertextual links must also be maintained. Should, for instance, a scholarly portal[8] that is no longer updated be preserved if it links to texts in many different places in many different countries, and if so, who should preserve it? As a curious anecdote one such rather large scholarly portal has completely vanished from the Web between the time of writing the original version of this paper and today.[9]

Interaction and Navigation

If we direct our attention to differences in possibilities for *interaction*, we can observe corresponding displacements. We navigate in digital collections via the interface of the computer, which is most often via interaction with symbolic representations on a screen. Paradoxically, the many spatial metaphors we use about the Internet (cyberspace, homepage, web site, navigation), reflect a relative movement from the archive as *place* to the archive as an *abstract textual zone.* The textual representation is, of course, in itself spatial, but in this respect it does not differ from earlier media of writing. As Pierre Lévy has observed, we do not move around in digital archives with books under our arms, but let the texts "move" on the screen (Lévy 1998, 57-58). Texts can now be made active in a very concrete sense: they can execute an embedded functionality, which we know from a link, or they can be made the object of automatic processing such as, for example, a query search. The active character of the text brings us back to the concept of hypertext, and a closer characteristic of this phenomenon on the basis of the active link.

Nelson's opposition of the linear printed text and non-linear hypertext has inspired a number of interpretations of the significance of hypertext, and the opposition is often extended by opposing the printed, fixed text and the interactive, dynamic digital text. In his *Writing Space*

from 1991, Jay Bolter argues that hypertext marks a radical break with "print culture."

> *The computer is calling into question the idea of fixity: in place of the stable printed text, the computer offers us a fluid, interactive text. The computer promises therefore to reverse the qualities that Eisenstein identified in the printing revolution.* (Bolter 1991, 22)

George Landow proposes a variant of this view of hypertext: "Hypertext fragments, disperses or atomizes text [...] by removing the linearity of print [and it] destroys the notion of a fixed unitary text" (Landow 1997, 64). Even though these visions can in principle be realized with hypertext, they express a certain vision for the use of hypertext more than a theory that describes the difference between printed and digital texts. Similar arguments are found in the writings of Richard Lanham (1993).

As a general theory about hypertext, the opposition of non-linearity and linearity is problematic in that we are familiar with many examples of books not intended for linear reading (Aarseth 1995). Reference works are the most obvious example, often containing highly complex internal cross-referential relations. If we shift focus from the individual text to collections of texts, such as libraries, non-linearity has always been the norm, and here the opposition appears completely absurd. In addition, the idea of the disintegration of the stable text is misleading if it is conceived as a general hypothesis. Even though digital texts can be dynamic and interactive, the practice we can observe on, among other places, the Internet suggests that many resources are being used precisely to guarantee that many kinds of texts are transmitted in an unaltered form. This is especially true for scholarly publication, whose entire modern system of footnotes and references builds on the validity of what they refer to – not as authoritative, but as actually existing and accessible, sources. Rather than defining hypertext as a particular non-linear narrative form, we can consider hypertext as characterized by the hypertextual link: *a rhetorical device characterized by the possibility of a user/reader-activated mechanized jump from one node to another based on the use of electricity* (Dalgaard 2001). Hypertext is thus neither a medium nor a special multi-linear form, but a rhetorical device among many

others in the computer. As an optional rhetorical device, it can be used for highly different ways of organizing information (classification systems, indexes and cross-references), and make use of highly different forms of expression (text, images, sound) for highly different purposes. Hyperlinks are also used in everything from non-linear fiction to references between texts and hierarchical classification systems. A clear tendency that sometimes drowns in the arguments about a new revolutionizing non-linear textuality and the fragmentation of the text in a loose network of text fragments, is that hyperlinks on the Internet have achieved their most pervasive significance at a level that does not concern the argument in the individual text, but the *collection* or *archive* of texts.

Hyperlinks support a predominant navigational mode on the Internet often described as *browsing*, but this is not the only way to navigate. Another predominant mode for navigating is the query, which is familiar from the digital catalogues of most libraries. Whereas browsing can be described as navigating through explicit choices, the query is characterized by the user contributing to the process with a combination of signs and starting an automated search process according to this combination. These modes are logically, functionally and practically different, but although at any given point in time we may use one modus or the other, navigation often occurs as a continual alternation between them. If one uses, for instance, a search engine like *Google* <http://www.google.com>, the result is often a list of links that allow the users themselves to continue browsing or carry out a more specific query within the results. In summary, the Internet gives texts and archives an altered relation to time and space and an altered potential for interaction, which opens up new possibilities and raises new problems. To understand how these are handled in practice requires a closer look at scholarly texts and archives on the Web.

Scholarly Texts and Archives on the Web

After starting out fairly tentatively at the beginning of the 1990s, the amount of digital texts, journals, and archives has grown dramatically, especially after the broad success of the World Wide Web in the middle of the 1990s. It is to a great extent journal literature, to a certain extent dissertations, and to a lesser extent published monographs that have

found their way to the Internet. It has been estimated that the number of "peer reviewed" journals on the Internet has increased from 27 in 1991 to around 8000 in 1998 (Harnad 1998), and a reasonable guess would be that there are 10,000-12,000 journals on the Internet today.[10] Most of the major publishers of journals now publish many journals in both printed and digital versions. In addition to this there are new forms of publications such as "e-print" archives, researchers who "publish" their text on personal or institutional web sites, and a profusion of specialized research archives. Even though the number of printed journals (still) far exceeds the number of journals on the Internet, there is thus already a substantial amount of scholarly literature on the Internet. It is not the aim of this article to reiterate the discussion of the extent to which digital forms of publications replace printed media and the possible timeframe of such an eventuality, but merely to point out that the Internet already constitutes an essential resource for access to scholarly texts. In the following we shall concentrate on two examples of web archives and their use of hyperlinks and search facilities in organizing scholarly texts.

The ACM Archive

ACM (The Association for Computing Machinery) is a scholarly society with around 80,000 members. It publishes a number of journals and "proceedings" from conferences organized by the many interest groups of the association. The digital archive was launched in 1997 and contains all ACM's publications dating back to 1991 and selected publications further back. In 2001, a search indicated a total of around 57,000 texts, a number that has grown to around 140,000 in 2004 (accessed 15 June 2004). In 2001, ACM furthermore supplemented their digital archive with a more general "Guide" for computing and computer-related literature, so now one finds an interface that makes navigation possible either in its aggregate bibliographical material of more than 750,000 citations or in ACM's more limited archive, which offers full-text access (accessed 15 June 2004). ACM's digital archive is among the most highly developed and integrated scholarly digital archives on the Web and constitutes an excellent example of the movement from publisher of journals to supplier of hypertextual archives on the Internet. The archive thus offers a good example for studying the significance of the Internet

for ACM's role as a publisher, for discussing navigation in the scholarly corpus, and for more generally looking at the textual reconfigurations attached to the hypertextualization of relations between texts.

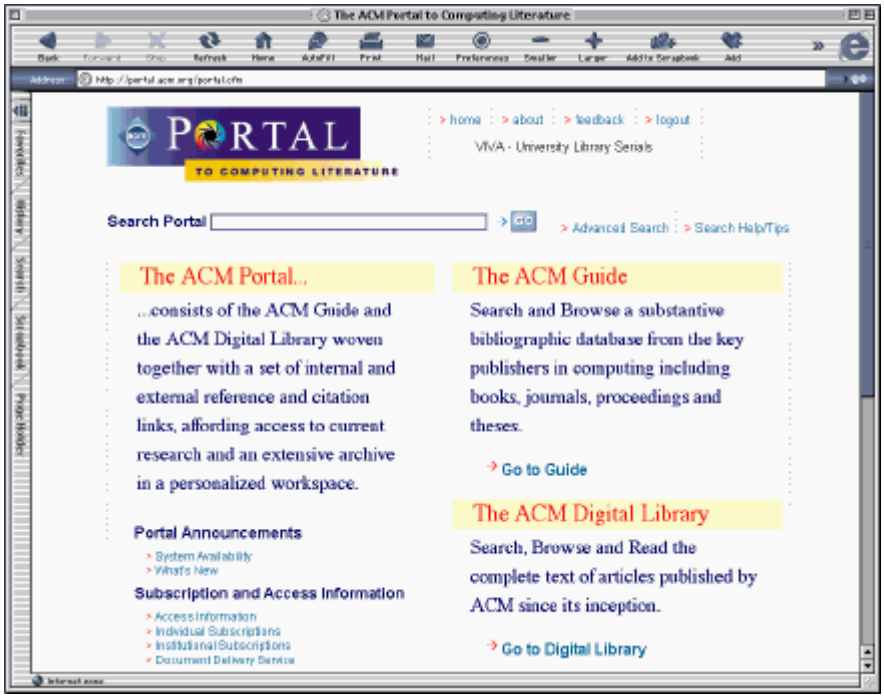

Figure 1. Excerpt from the ACM portal as it looked on November 30, 2001. This entrance web page gives access to either its aggregate bibliographical resources, ACM's full-text archive, or both.

By creating a hypertextual archive ACM has in many ways expanded its role in the scholarly system of knowledge. Its role as a publisher of journals and conference publications is now supplemented by the organization of these publications in an archive serving as a global library for this literature. The individual publication (e.g., journal or conference transaction) is thus embedded in a greater context in which, for instance, the title of the journal is only one among many organizational principles supported in the ACM archive. The latest addition, the Guide to all of

the literature within a subject area, marks a further extension of the scholarly breadth and depth of ACM's profile, not as a publisher of texts, but in a classificatory and navigational respect. The Guide as a bibliographic collection aiming for universal coverage of the field of computing illustrates and remedies the somewhat arbitrary limits of the ACM full-text collection, which, as other publisher-archives, holds only a fraction of the literature within its field(s). Since it is possible to follow references via links, automatically obtain other articles by the same author, or see related articles with overlapping keywords, it does not make any sense to let these possibilities be arbitrarily limited by whether or not ACM is also the publisher of this related literature. One of the criteria for inclusion in the Guide is, notably, if the text in question is referenced by an ACM publication – or in other words, if there is basis for a link.

ACM's archive can be browsed according to a wide range of different criteria such as publication, author, title, subject and keyword, or one can search using these criteria, as well as search for words in the abstracts and the full text and browse further from the results of this kind of query. Navigating in an archive like this is thus based on a bunch of different textual forms that constitute the context and interface for the primary texts in the archive. We can divide these different texts into three main types based on their relation to the primary texts. *Paratexts* are the texts that comprise the individual text's interface such as its title, author, abstract and year, where the author/publisher is the sender. *Intertexts* are the explicit references between primary texts.[11] Finally, *metatexts* are the subject and classificatory texts used to position the primary texts within the individual archive. ACM has for instance developed its own classification system and subject index in which each primary text is grouped.

Neither the paratexts, intertexts and metatexts nor the issues surrounding the organization of texts are new in and of themselves. But in hypertextual archives like ACM a number of displacements nevertheless seem to be taking place in the interface between text and archive established by paratexts, intertexts, and metatexts.

The most immediate difference is, of course, that the cross-referential relations we know from catalogues, indexes and references are now built into the archive as hypertextual links – they shift from being purely rep-

resentational entities to also being active mechanisms. Another, and related, consequence is that these well-known intertextual, paratextual and metatextual forms can – and are – configured in new constellations. When one has navigated to the bibliographic entry for a text in the ACM archive, one meets a web page composed of a number of the above-mentioned textual forms.

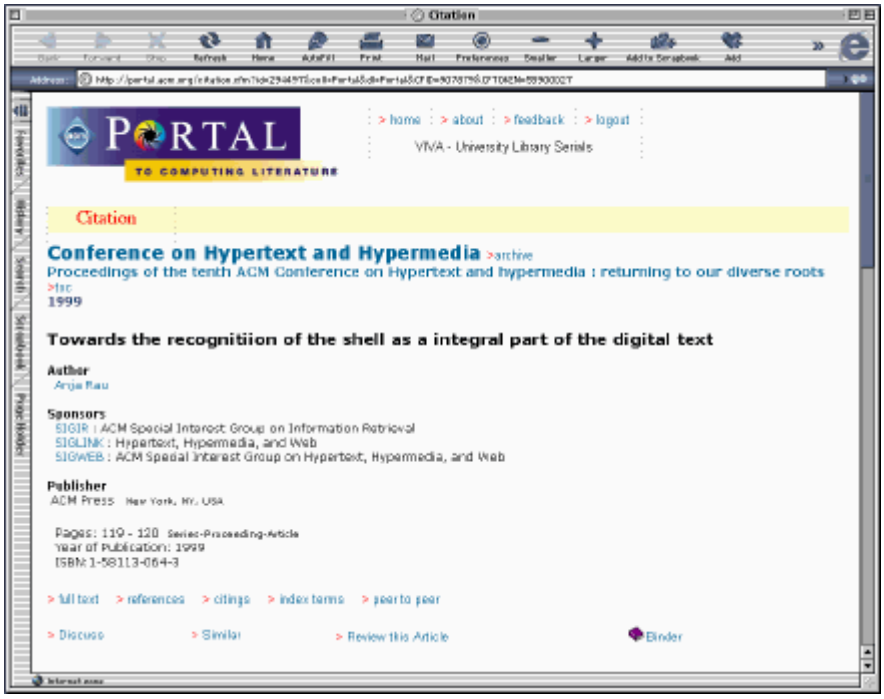

Figure 2. Excerpt from a bibliographic entry in the ACM Digital Library (accessed 30 Nov. 2001). Below the screen shown here are, on the same web page, references, an index, the location of the text in the classification system and a link to the represented text itself, among other things.

In addition to a link to the text itself, this page contains the following: bibliographic information such as author, date, and publisher (paratext); the references (intertext) of the article in question; its primary

classification and secondary classifications as well as keywords and index concepts (metatext). The bibliographic entry also contains intertextual forms known from abstract and indexing services, most notably the inclusion in the entry of citations from later published literature. Unlike print-based abstract services, which are detached from the actual literature, digital collections such as the ACM can provide hyperlinks to such subsequent works that cite a given work. In this way the text is positioned in a new way in the scholarly literature – it no longer only contains references to its own development, but also to its reception and use, and more importantly it offers access in both directions. Another innovation consists in the possibility for automatically generating a list of "related articles," based on overlapping references and keywords. Admittedly, the usefulness of the current implementation of this sorting mechanism in the ACM archive is not convincing, but it reflects an attempt to textualize potentially relevant relations between texts not traditionally made explicit – not unlike the situation when a colleague suggests a text based more or less on knowledge of one's field of interest.[12] A third example of innovation is the presence of "peer-to-peer" articles, meaning that the reader is provided with a recommendation indicating that other readers who have read (in fact only downloaded) article *x* have also read articles *y* and *z* that may be of interest to the present reader.[13] What is new in these examples is, of course, neither the existence of related articles nor that articles cite other articles. The change consists in these relations, which already exist explicitly or implicitly, being represented textually in the bibliographic entry and built into the archive as active links.

We can thus consider the way in which hyperlinks change the way we navigate between these textual forms. Different organizational forms like classification systems, indexes and networks of references support different "reading modalities;" one can move up or down through a classification tree, while references operate across primary texts and establish a history between them. If we consider the complex textuality of the bibliographic text as a navigational tool, it is striking that one can "move" in almost any direction from here. One can for instance move vertically up or down in the classification tree used by ACM, or one can

move horizontally via the network of references to previous or later texts that the bibliographic entry has "extracted" from the primary texts.

The technical integration of these forms of organization in the ACM archive indicates that the ability to switch flexibly between them and the associated "reading modalities" constitutes an important new feature in the navigational framework that the digital archive can offer. This kind of rapid and flexible navigation has hitherto characterized individual texts and very small collections, but now becomes a feature also of the archive, which can be navigated to an extent and with a speed never attainable in a traditional library. This does not, however, make the archive simpler to use; the semantic relations established in an archive like the ACM archive themselves constitute a metatextual context that must be "read" for the navigational facilities to make sense in respect to one's own special focus. The way in which this context is established seems to be very different from archive to archive, which for one thing raises the question of the relation between these archives on the Internet and how to navigate between them. This issue will be taken up subsequently, but first we shall take a look at a research archive that distinguishes itself from the ACM archive as concerns its scholarly content, editorial profile and organization of resources.

The Rossetti Archive

The Rossetti Archive was developed at the University of Virginia's Institute for Advanced Technology in the Humanities, where around forty major and minor projects of this kind have been developed. This archive is interesting because in many respects it is a theoretically motivated attempt to rethink a classic humanistic textual form, the critical edition. The main architect behind the project, Jerome McGann, has thus formulated his visions for a new kind of text-critical version both before and during the development of the Rossetti Archive (McGann 1991, 1992, 2001).

The Rossetti archive has been under development since 1993, and in 2000 a temporary version was published as a web archive; the first official version of the archive was introduced in 2002. The archive contains Dante Gabriel Rossetti's (1828-1882) collected works, both texts and images, in their various forms and versions (manuscripts, published texts,

paintings, photographs of paintings). In addition to this are extensive bibliographic material, introductory comments, contemporary related works and publications, a Rossetti biography and a link collection to more recent articles about the archive and Rossetti. The archive currently consists of around 10,000 image and text files and is expected to reach double this amount in its final version.

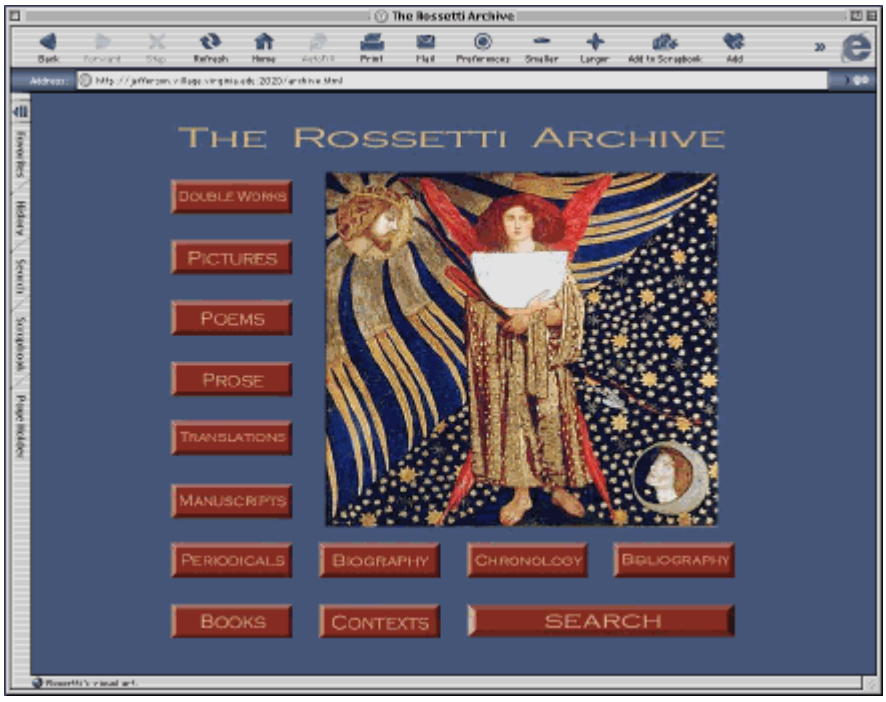

Figure 3. Image of the central web page in the Rossetti archive (accessed 30 Nov. 2001).

The Rossetti Archive has grown out of the meeting between a new medium (the computer), a concept of the text and a critique of predominant text-critical traditions that McGann started developing in the beginning of the 1980s. The idea has been to develop a model for hypermedia archives that can comprise two different approaches to scholarly edited texts: the text-critical version (with focus on the histori-

cal relations between various, often lacking or altered, versions of a text) and the "exact" version (with focus on the most exact rendering of a particular text in the sense of physical artefact, often rare works) (McGann 2001, 11). The first and predominant approach thus emphasizes what McGann labels the linguistic code of a text whereas the latter approach gives priority to the materiality and visuality of the text, labeled the bibliographic code of the text. The purpose of bringing these edition types together is rooted in McGann's theoretical perspective on textuality, according to which the text's material and visual form is regarded as being just as significant as the language. Another one of McGann's fundamental ideas is that text editing must deal with the fact that texts are produced, exist and are received in a social context, or, as he puts it, "that the social intercourse of texts – the context of their relations – must be considered an essential part of the 'text itself' if one means to gain an adequate critical grasp of the textual condition" (McGann 2001, 12). This means that different editions and versions of a particular text are not necessarily edited into "the original" text, based on an ideal about approaching the author's original intentions, but are instead considered equally valid versions of a particular work. This position multiplies scholarly editing, or in any case the text in which it results, because the single, critical, authoritative text is replaced by several parallel equivalent versions or texts.

Because of their character, Rossetti's works offer an suitable material and a serious challenge for these ideas. Rossetti worked not just as a writer and painter, his texts were often interpretations of motifs from his paintings, and in some cases his drawings and paintings were a treatment of the texts he had written. He rewrote previously published texts and treated the same motif in many images, which is why his works are available in many variants and in different constellations. Rossetti's works therefore constitute a complex corpus with many relations that for McGann constitute an aggregate whole that scholarly editing must necessarily include if it is to form the basis for criticism and interpretation of Rossetti: "[…] these features of Rossetti's work pose a complex and hitherto unsolvable editorial problem. One cannot properly study or appreciate Rossetti's work without having access to all of it" (McGann 2001, 13). This requires, claims McGann, a technology that is more

complex than the book – a hypermedia archive that does not merely contain the collected works, but is also capable of rendering visible relations between the works, placing them in relation to one another and establishing a context for understanding them. The idea is thus to create what may be called an explorative archive for Rossetti's collected works.

The Rossetti Archive is organized according to three main groups of resources: a genre classification of Rossetti's works (poems, images, prose, translations and double works), a material collection (books, manuscripts, journals) and a grouping of related resources (bibliography, biography and related works by others). The construction of the archive is thus uniquely related to Rossetti's artistic production, which the introduction of the category "double works" suggests. The concept of double works refers to the instances where Rossetti treated the same theme in both texts and images, such as "The Blessed Damozel." If one browses or searches for this work, the sonnet and the painting are brought together, and in the final version of the archive there will be access to the more than forty versions of the sonnet that have been published and the almost twenty drawings and paintings available. Each of these resources is considered a "Rossetti Archive Work" (RAW) and accompanied by source-critical information and commentary that place the individual RAW in a context. These are kept together by a more general concept of the work, represented in the shape of a "central commentaries page," which places a work like "The Blessed Damozel" in a larger context through detailed commentaries on its creation and artistic significance.

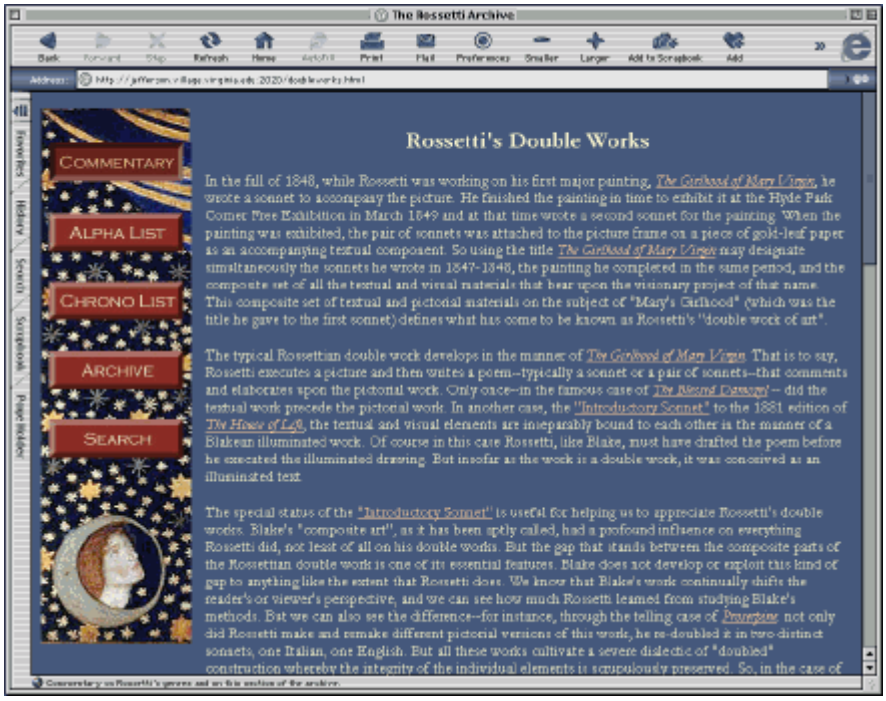

Figure 4. Excerpt from a web page of the Rossetti Archive under the category "double works" (accessed 30 Nov. 2001).

A number of technical and editorial issues are, of course, connected with digitally representing texts and images, organizing these resources and developing an interface that makes possible structured queries and browsing among them. However, here we will concentrate on the shift or reconfiguration implied by the editing of an archive in comparison to the editing of a book. One should be careful about contrasting the Rossetti Archive with a typical critical version in book form, as the archive does not constitute an individual text, but is a series of texts (works) and establishes a metatextual context for these. But it is precisely this difference that holds the key to understanding the groundbreaking aspects of the Rossetti Archive: a shift in the editorial perspective from the individual archive to the collection of works, and from the individual

work in the Rossetti Archive as unique entity to the work as a plurality of versions that are openly presented. Viewed separately, most of the works in the Rossetti Archive would certainly be preferable in book or picture form, but as an aggregate corpus the archive offers something else: the possibility of juxtaposing works, of comparing versions, of orienting oneself through the enormous collection of source-critical information and commentaries accompanying each main group, each work, and each version of a particular work. An archive with many texts does not put an end to the editorial issues for the individual text; rather it adds an extra editorial dimension concerning the editing of the archive as such. It is not merely the individual text that must be edited, but also its relations to other texts and works in the archive. The editor and the user/reader of this kind of archive thus operates with a metatextual context that is not merely established with respect to the individual work, but with respect to the collected corpus of works and variations of these works and the contexts of their publications.

The full yield of an archive like the Rossetti Archive requires an interest in and knowledge of Rossetti that only a limited number of scholars and interested parties have, and we will leave the evaluation of its scholarly value to them. But it seems reasonable to conclude that the archive is an example of an attempt to rethink the scholarly edition as a scholarly archive that encompasses a host of material, which is practically outside the scope of any print edition but still retains an editorial structure and organization very different from a raw collection of materials.

The individual work or versions of a work is still a pivotal point, but the intertextual relations between works and the metatextual commentaries positioning it within the framework of Rossetti's entire artistic production establishes a context that brings forth its relations to surrounding works. Thus there is an editorial perspective behind it and a series of interpretative commentaries that point to relations between materials, which in some respects illustrate a change of focus from the individual authoritative edition to the comprehensive and edited archive.

In spite of their obvious differences, or perhaps precisely because of these, there is reason to take note of a feature shared by the ACM Archive and the Rossetti Archive. The use of hyperlinks and search facilities does not seem to greatly influence the primary texts. It is to a

greater extent in the interface between text and archive – in the textual apparatus where paratexts, intertexts and metatexts operate – that we see the significance of hyperlinks. If hyperlinks promise to change the way we can navigate in our corpus of knowledge, it is not because they are used to establish radically new relations between texts, or, as Bolter and Landow seem to imagine, to break radically with the text as a communicative form, but rather because hyperlinks "realize" and support a series of relations already existing between texts. However, this is in itself a change that has the character of a *realization* of prior *virtual* relations – a movement that goes against the typical description of the Internet as virtual.

Another shared feature is the move from publishing an individual text/publication to publishing and organizing a total corpus, or in other words, a shift from text to metatextual context. In the ACM Archive, the individual journal and the individual conference publication are embedded in the aggregate archive as the unit to which one subscribes and has access. Correspondingly, the Rossetti Archive is published as an aggregate integrated archive and not as a series of isolated works.

The Web as an Archive of Archives: Network Logic and Complexity

Up until now we have considered the ACM Archive and the Rossetti Archive as isolated archives, but they exist, of course, in the broader context comprised by the Internet. If we consider the World Wide Web as one big system, it is a highly anarchistic network to which everyone with access to the Internet can contribute (and many do), but without any coherent organization or guarantee of the durability of links and resources. This also applies to the ACM Archive and the Rossetti Archive in the instances when they have links to resources outside of the archives themselves.

The Web manifests the ideal of a global archive in many ways, but not in the sense of a well-ordered, cataloged and well-planned system of knowledge. The Web is thus an archive in the sense that, in principle, one has access to everything, but not an archive in the sense that there is an institution in charge of organizing and preserving its contents. Even if we focus on scholarly texts, which constitute only one among many

communicative genres on the Internet, there is no distinct structure, but rather a loose network that is not organized according to some shared higher principle. This does not mean that there is no order or structure on the Internet; there is a lot of structure, hierarchy and control, but of a more or less local nature, as is the case with the ACM Archive and the Rossetti Archive. In this broader context, these archives, and the many others that exist, constitute different, often partially overlapping, editorial excerpts from the total, loosely connected corpus of scholarly texts on the Internet.

This is one of the reasons why the concept of the network at this level, in spite of its dispersed and often diffuse use, reflects an actual tendency in principles underlying the organization of texts. It is a widespread practice to oppose the network structure with the tree structure (Bolter 1991; Deleuze & Guattari 1987; Berners-Lee & Fischetti 1999). This is an understandable opposition for both logical and historical reasons; logically because the tree structure differentiates itself from the network structure by the kind of relations that exist between its units, and historically because the tree structure has been the predominant metaphor and the predominant principle for organizing knowledge in the typographical system of knowledge. The concept of the "tree of knowledge" was thus both a widespread metaphor and an objective in the period after the printing press (Burke 2000).

There are several reasons to exercise caution in transferring this opposition to an opposition between digital and typographical organization of knowledge. First, long before the Internet existed there were other principles for organizing knowledge besides the tree structure; for instance, alphabetic ordering, in itself meaningless, gained ground concurrently with the growth of knowledge after the printing press and brought about the need to embed a network of cross-references in reference works (Burke 2000). In the same way the modern system of references – an intertextual network between primary texts – can be viewed as a product of the stabilization of the individual version that the printing press brought about. Thus, both network structures and a flat neutral structure such as alphabetic organization have thrived with print. Finally, the Internet is filled with tree structures (e.g. classification systems) that are not on the verge of extinction but on the contrary are

also augmented by hyperlinks. The ACM Archive, for example, testifies to the fact that both structures are alive and well as organizational and navigational principles. The relation between network and tree structure has changed with the introduction of a new medium, but it is not that of a mutually exclusive antagonism. The change can tentatively be illustrated by the observation that the network previously existed within the framework of the "Tree of Knowledge," while the tree structure on the Internet exists as local orders in a major network. The success of the concept of network is an expression of the breakdown of the idea that we can organize our total corpus of knowledge in a coherent and meaningful system, not the extinction of hierarchies. In one sense the Web as an archive of archives is a manifestation and a model of a pluralistic system of knowledge as opposed to a unified one – in this respect, however, it reflects not a break but a continuation of tendencies identified already by Wells and Bush.

Another feature that distinguishes the Internet as a meta-archive is complexity. The technical integration that takes place in digital archives, where one can move directly between texts that refer to each other and search in a variety of ways in and among archives on the Internet, is sometimes described as seamless and intuitive. It is therefore essential to point out that this movement is not countered by a semantic integration out of which a simpler and uncomplicated archive grows, on the contrary: the complexity increases not only in the Rossetti Archive, which aims at critically and reflectively using and investigating relations and meaning in and between Rossetti's works, but also in the ACM Archive, where the internal organization establishes a complex system of relations and organizational principles that position the text in a variety of ways. This tendency becomes even clearer if we consider the Internet more generally. It is a question of growing semantic complexity, attested to by the increasing amount of overlapping archives, portals and personal link collections on the Internet.

The tendency to expand the textual apparatus surrounding the primary texts is not without parallels to the period after the printing press, when classification systems and reference works were reformed, as Eisenstein (1979) and others have described. Then, too, these efforts

constituted an attempt to organize and bring under control a tremendous growth in the amount of texts:

> *The printing industry had two important consequences in this domain [Burke is here referring to encyclopedias]. In the first place, it obviously made encyclopedias more readily and more widely available. In the second place, it made them even more necessary than they had been before the invention of the press. To be more precise, one of their functions became increasingly necessary, that of guiding readers through the ever-growing forest – not to say jungle – of printed knowledge.* (Burke 2000, 109)

The possibility of directly accessing this enormous corpus of texts on the Internet and the possibility of moving in and out the text collections of various scholarly societies alleviates the researcher's access to texts, but at the same time increases the demands on competent navigation. If one chooses to make use of one or a few archives or portals, it is with an awareness of the existence of many other potentially relevant knowledge filters with other editorial criteria and principles for organizing and navigating. While the prospects of a final solution for the organization of knowledge seem to be fading concurrently with the growth of the Internet, there is nothing to prohibit increased reflection on navigation and the organization of knowledge. The development of critical competencies in this area is imperative if we wish to find the narrow path between being overwhelmed by and ignoring the amount of information on the Internet.

Rune Dalgaard is a lecturer at the Institute for Information and Media Studies, Aarhus University, Denmark. He has earned a MA degree from the same institution and is currently pursuing a Ph.D. degree. More information and publications on topics such as media theory, hypertext, media history and technology studies can be found at <http://www.imv.au.dk/medarbejdere/runed>.
E-mail: runed@imv.au.dk

Notes

1. This article is a translated (tr. Stacey Cozart) and slightly revised version of "Videnskab.net" in *Cyberkulturer og rekonfigurationer*, ed. Mette Bryld & Randi Markussen, København: Samfundslitteratur, 2003, 211-239. Most of the empirical work dates back to 2001, but the main observations have not been affected by later developments.
2. The concept remediation has been introduced by Jay Bolter and Richard Grusin (1998) and directs attention to the relation of sameness and difference in the evolution of media forms, the multiple ways in which digital media remediate earlier visual media forms, and the two-way direction of remediation as old media also borrow from new ones. It is noteworthy, also, because it marks a shift in the thinking of Bolter, who in earlier works, most notably *Writing Space* (1991), claimed that digital text was a radical break with print culture and would replace it. The basic idea that new media remediate old media is well known from the media theoretic works of Harold Innis, Marshall McLuhan, Eric Havelock, Elizabeth Eisenstein, Jack Goody, Joshua Meyrowitz and many others. Bolter and Grusin contribute most notably the idea that remediation is characterized by the two reverse trends of immediacy (striving for transparent mediation) and hypermediacy (striving for visibility of the medium and mediation).
3. The issue was analyzed in greatest detail by Derek de Solla Price in the 1960s and 1970s (see, for instance, Solla Price 1965). In Denmark, Johan Fjord Jensen wrote a brilliant critical essay on the relationship between the flood of information and research control in the book *Babel og Tomrum* (1996). Among other works, Jensen draws on that of Solla Price.
4. Matthew Kirschenbaum (2003) in an unpublished excerpt from his forthcoming book has suggested that the first digital library was a system called Professor Ramac from 1958 based on the then new random-access storage technology, the hard disk. See <http://www.otal.umd.edu/~mgk/blog/archives/000237.html>.
5. For literature on the development of the Internet, see Abbate 1999 and Hafner & Lyon 1996.
6. Several sources date the first use of the concept of hypertext to 1965. Nelson himself claims to have published the term hypertext in 1965 (Nelson 1987, 1/29). According to some accounts he first used the concept of hyper-text during a lecture at Vassar College, New York. See Kirschenbaum 2000, 122.

7. Based on empirical observations, John Morkes and Jakob Nielsen (1997) reached the conclusion that people read 25% slower on screen than on paper and that long texts were seldom read on the screen. A classic, and often mentioned, example of a defense of the book based not only on "usability" criteria but on moral and aesthetic grounds is Sven Birkerts' *The Gutenberg Elegies: The Fate of Reading in an Electronic Age* (1994). This work can in part be viewed as a reaction to a series of hypertext theories in the early 1990s, most notably Bolter (1991), Landow (1992) and Lanham (1993).
8. The phrase "scholarly portal" refers to an editorially organized collection of references to scholarly resources within a subject area that gives access to these resources via hyperlinks. See, for instance, *Netzwissenschaft.de/Net.science* <http://www.netzwissenschaft.de/e.html> (although cf. next note), an interdisciplinary portal for Internet research; *Voice of the Shuttle* <http://vos.ucsb.edu/>, a humanities portal; or, for the sake of contrast, *PubMed* <http://www.ncbi.nlm.nih.gov/entrez/>, a portal in the area of biomedicine. Scholarly portals are closely related to subject indexes, scholarly bibliographies and abstract services, which are all familiar from the typographical system of knowledge, but on the Internet usually offer direct hyperlinks and thereby access to the resources they index.
9. The portal in question is *Netzwissenschaft.de* (English version: *Net.science*) that disappeared from one day to the next. E-mailing the editor, Reinhold Gether, has not produced any response.
10. To my knowledge, an authoritative estimate of the number does not exist. In a letter to the author, Ann Okerson, an authority in the area and editor of *NewJour* (an index of journals and mailing lists on the Internet, <http://gort.ucsd.edu/newjour/>), stated that *NewJour* did not link to all existing journals and estimated that we are approaching 10,000 journals (13 Aug. 2001).
11. These concepts are inspired by Gérard Genette's theory of "transtextuality," where he operates with intertextuality, paratextuality, hypertext, metatext and architext. Genette's shortest definition of the concept of transtextuality reads: "the text interests me (only) in its *textual transcendence* – namely, everything that brings it into relation (manifest or hidden) with other texts. I call that *transtextuality*" (Genette 1992, 81). For further discussion of Genette's concepts and my use of them, see Dalgaard 2001.
12. A better implementation, in my own view, can be found in the fully automated digital collection *CiteSeer*, which automatically indexes scientific literature on the Web, see <http://citeseer.nj.nec.com/cs>.
13. This feature, technically known as "collaborative filtering," has been made popular by *Amazon*, where it also works significantly better than in the ACM collection. See <http://www.amazon.com>.

References

AARSETH, ESPEN J. (1995). *Cybertext: Perspectives on Ergodic Literature*. Diss. Bergen: University of Bergen.

ABBATE, JANET (1999). *Inventing the Internet.* (Inside Technology). Cambridge, MA: MIT Press.

The ACM Portal (2004). The Association for Computing Machinery. <http://portal.acm.org/portal.cfm> [2001]

BENIGER, JAMES R. (1986). *The Control Revolution: Technological and Economic Origins of the Information Society*. Cambridge, MA & London: Harvard UP.

BERNERS-LEE, TIM & MARK FISCHETTI (1999). *Weaving the Web: The Past, Present and Future of the World Wide Web*. London: Orion Business.

BIRDSALL, WILLIAM F. (1994). *The Myth of the Electronic Library: Librarianship and Social Change in America*. Westport, CT: Greenwood Press.

BIRKERTS, SVEN (1994). *The Gutenberg Elegies: The Fate of Reading in an Electronic Age*. Boston: Faber and Faber.

BOLTER, JAY D. (1991). *Writing Space: The Computer, Hypertext, and the History of Writing*. Hillsdale: Lawrence Erlbaum.

BOLTER, JAY D. & RICHARD GRUSIN (1998). *Remediation: Understanding New Media*. Cambridge, MA: MIT Press.

BURKE, PETER (2000). *A Social History of Knowledge: From Gutenberg to Diderot*. Malden, MA: Polity.

BUSH, VANNEVAR (1991/1945). "As We May Think." *From Memex to Hypertext: Vannevar Bush and the Mind's Machine*. Eds. James Nyce & Paul Kahn. San Diego & London: Academic Press. 85-110.

CiteSeer.IST: Scientific Literature Digital Library. <http://citeseer.nj.nec.com/cs> [2004-06-30]

DALGAARD, RUNE (1999). "Mediesfærer: om mediers betydning for overlevering af viden." Master's thesis. Aarhus: University of Aarhus. <http://imv.au.dk/~runed/pub/pub.html> [2004-07-06]

DALGAARD, RUNE (2001). "Hypertext and the Scholarly Archive: Paratexts, Intertexts and Metatexts at Work." *Proceedings of the 12th ACM Conference on Hypertext and Hypermedia, 14-18 August, 2001*. Aarhus: Association of Computing Machinery. 175-184. <http://imv.au.dk/~runed/pub/pub.html> [2004-07-06]

DELEUZE, GILLES & FÉLIX GUATTARI (1987). *A Thousand Plateaus: Capital and Schizophrenia*. Minneapolis: University of Minnesota Press.

EISENSTEIN, ELIZABETH L. (1979). *The Printing Press as an Agent of Change*. 2 vols. Cambridge: Cambridge UP.

ENGELBART, DOUGLAS C. (1963). "A Conceptual Framework for the Augmentation of Man's Intellect." *Vistas in Information Handling.* Eds. P.W. Howerton & D.C. Weeks. Vol. 1. London: Spartan. 1-29.

ENGELBART, DOUGLAS C. (1984). "Authorship Provisions in Augment." *Intellectual Leverage: The Driving Technologies: Compcon 84, 28th IEEE Computer Society International Conference*. San Francisco: IEEE Computer Society Press. 465-472.

FINNEMANN, NIELS O. (1994). *Tanke, Sprog og Maskine: en teoretisk analyse af computerens symbolske egenskaber*. København: Akademisk Forlag.

FINNEMANN, NIELS O. (1998). "Computeren: et medie for en ny skriftteknologisk revolution." *Multimedier, Hypermedier, Interaktive Medier*. Ed. Jens F. Jensen. Ålborg: Ålborg Universitetsforlag. 43-68.

GENETTE, GÉRARD (1992). *The Architext: An Introduction*. Trans. J.E. Lewin. Berkeley, CA: University of California Press.

HAFNER, KATIE & MATTHEW LYON (1996). *Where Wizards Stay Up Late: The Origins of the Internet*. New York: Simon & Schuster.

HARNAD, STEVAN (1998). "On-Line Journals and Financial Fire-Walls." *Nature* 395 (10 September): 127-128. <http://www.cogsci.soton.ac.uk/~harnad/nature.html> [2001-12-14]

JENSEN, JOHAN F. (1996). *Babel og Tomrum: de systemiske videnskaber og humaniora*. København: Gyldendal.

KIRSCHENBAUM, MATTHEW G. (2000). "Hypertext." *Unspun: Key Concepts for Understanding the World Wide Web*. Ed. Thomas Swiss. New York: New York UP. 120-137.

KIRSCHENBAUM, MATTHEW G. (2003). "An Excerpt from *Mechanisms* [2]: 'Professor Ramac'." <http://www.otal.umd.edu/~mgk/blog/archives/000237.html> [2004-01-15]

LANDOW, GEORGE P. (1992). *Hypertext: The Convergence of Contemporary Critical Theory and Technology*. Baltimore: Johns Hopkins UP.

LANDOW, GEORGE P. (1997). *Hypertext 2.0: The Convergence of Contemporary Critical Theory and Technology*. 2nd rev. ed. Baltimore: Johns Hopkins UP.

LANHAM, RICHARD A. (1993). *The Electronic Word: Democracy, Technology, and the Arts*. Chicago: University of Chicago Press.

LÉVY, PIERRE (1998). *Becoming Virtual: Reality in the Digital Age*. New York: Plenum.

LICKLIDER, JAMES R. (1965). *Libraries of the Future*. Cambridge, MA: MIT Press.

MCGANN, JEROME (1991). *The Textual Condition*. Princeton: Princeton UP.

MCGANN, JEROME (1992). *A Critique of Modern Textual Criticism*. 2nd ed. Charlottesville, VA: University Press of Virginia.

MCGANN, JEROME (2001). *Radiant Textuality: Literature after the World Wide Web*. New York: Palgrave.

MCLUHAN, MARSHALL (1962). *The Gutenberg Galaxy*. Toronto: University of Toronto Press.

MORKES, JOHN & JAKOB NIELSEN (1997). "Concise, SCANNABLE, and Objective: How to Write for the Web." *useit.com*. <http://www.useit.com/papers/webwriting/writing.html> [2004-06-30]

NELSON, THEODOR H. (1987/1981). *Literary Machines*. Swarthmore, PA: Published by author.

Netzwissenschaft.de/Net.science. <http://www.netzwissenschaft.de/e.html>

NewJour (1993-2003). Eds. Ann Shumelda Okerson & James J. O'Donnell <http://gort.ucsd.edu/newjour/> [2004-06-30]

NYCE, JAMES & PAUL KAHN, EDS. (1991). *From Memex to Hypertext: Vannevar Bush and the Mind's Machine*. San Diego & London: Academic Press.

OKERSON, ANN. Letter to the author. 13 Aug. 2001.

ONG, WALTER J. (1988). *Orality and Literacy: The Technologizing of the Word*. London: Routledge.

PubMed. National Center for Biotechnology Information. <http://www.ncbi.nlm.nih.gov/entrez/> [2004-06-30]

The Rossetti Archive (2000). Ed. Jerome J. McGann. Institute for Advanced Technology in the Humanities, University of Virginia. <http://www.iath.virginia.edu:2020/> [2001]

SOLLA PRICE, DEREK J. DE (1965). *Little Science, Big Science*. New York: Columbia UP.

WELLS, H.G. (1967/1936). "World Encyclopedia." *The Growth of Knowledge: Readings on Organization and Retrieval of Information.* Ed. Manfred Kochen. New York & London: Wiley. 11-22.

Voice of the Shuttle. Ed. Alan Liu. University of California, Santa Barbara. <http://vos.ucsb.edu> [2004-06-30]