MAX-PLANCK-GESELLSCHAFT

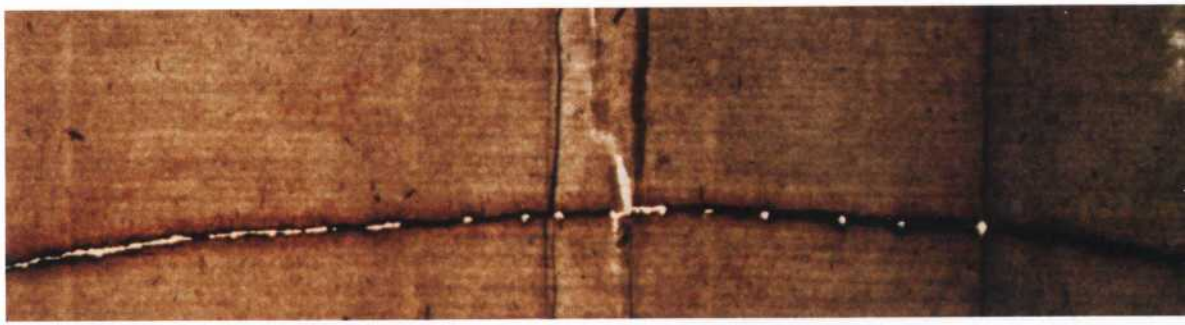
ECHO

EUROPEAN CULTURE HERITAGE ONLINE

Towards a Web of Culture and Science

Editorial

The rapidly increasing role of the Internet for a knowledge-based society represents a major challenge for cultural heritage and science as we know it today. The crisis of information management in the natural sciences is evident from ever-growing costs for scientific journals as well as from inadequate access structures which document the failure of transferring the established publication system of the natural sciences from the print to the electronic medium. Within the present, commercially dominated system of dissemination and access, science is simply unable to reveal its full impact so that investments in science fail to reach the returns they could in principle attain. Equally dramatically, cultural heritage as well as the humanities dedicated to its study are in danger of being left behind by technological development. The main body of sources constituting cultural heritage has in fact not yet been transferred to the new medium. This deficit concerns not only the humanities, which vitally depend on information representing cultural heritage, but also society at large.
The humanities and the social sciences so far have been incapable of initiating a dynamics comparable to that launched by the new



information technologies in the natural sciences. The imminent risk of culture and science losing ground in the new medium and the society shaped by it has been identified in several international strategy documents pointing to the lack of implementation of cultural resources in the new electronic media, as well as of sustainable standards and adequate tools. Existing programs have, however, so far failed to launch a self-sustaining infrastructure and dynamics for establishing a synergy between cultural heritage, scientific knowledge, and technological potential. If such a dynamics could be launched, on the other hand, it would massively challenge traditional boundaries blocking cooperation between different national cultures and scientific disciplines such as those related to lan-guage, thus creating entirely new potentials for research and public culture.

This report is dedicated to the challenges of the ongoing information revolution for culture and science. It gives an account of the recent efforts to create a future-oriented infrastructure on the Web ensuring open access to these fundamental resources of humanity, in particular the Berlin Declaration launched by the Max Planck Society and the European Cultural Heritage Online (ECHO) Initiative. Such an infrastructure should ease the transfer of contents relevant to culture and science from the traditional media to the medium of the future, it should exploit and further develop the technical potential inherent in the Web, and should turn the Web into an open, interactive, and sustainable public think-tank strengthening our capacity to solve the global problems of mankind. Readers of this booklet are invited to join the process initiated by the Berlin Declaration and the ECHO Initiative, subscribing to an open-access policy and enriching a growing infrastructure for scientific knowledge and cultural heritage giving rise to the Web of the future which will have to be a Web of Culture and Science.

"Cultural Heritage encompasses material culture,
in the form of objects, structures, sites and landscapes,
as well as living (or expressive) culture as evidenced
in forms such as music, crafts, performing arts,
literature, oral tradition and language.
The emphasis is on cultural continuity from the past,
through the present and into the future,
with the recognition that culture is organic and evolving."

The World Bank summary of a meeting in
Washington, 26-27 January 1988

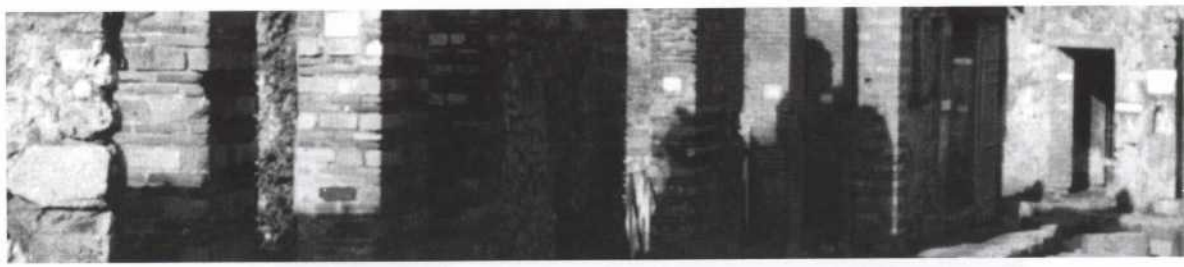
Chapter 1

The Crisis of Culture and Science in the Information Age

1.1 Challenges of the Information Revolution for Culture

This is a time in which technology and culture seem to be as decoupled from each other as they have ever been in the recent past. Technological visions of progress, in particular, have lost their appeal of being guarantors of the progress of culture into the bargain. While catchwords such as "information society" or "postgenomic society," not to speak of "traffic of the future," have lost glamour and credibility as promises of a better civil society, scepticism if not hostility with regard to science and technology are spreading. European culture, in particular, jointly created by the homo faber and the homme des lettres, faces a crisis: while it provided the foundation for magnificent technological achievements in a long-range development reaching back to antiquity, cultural heritage and its values are dramatically losing ground in the techno-scientific world that has emerged from it.

The medium of today and tomorrow, the Internet, might in fact leave behind a culture which is the heritage of our past but urgently needed to meet the challenges of the future. This cultural heritage, which binds us together even more strongly than our institutions, is presently in danger of being left behind, of missing the train, so to speak, of the rapid technological developments carrying us into a new information age. Moreover, wars and dwindling public funds for the preservation of cultural heritage are contributing to its

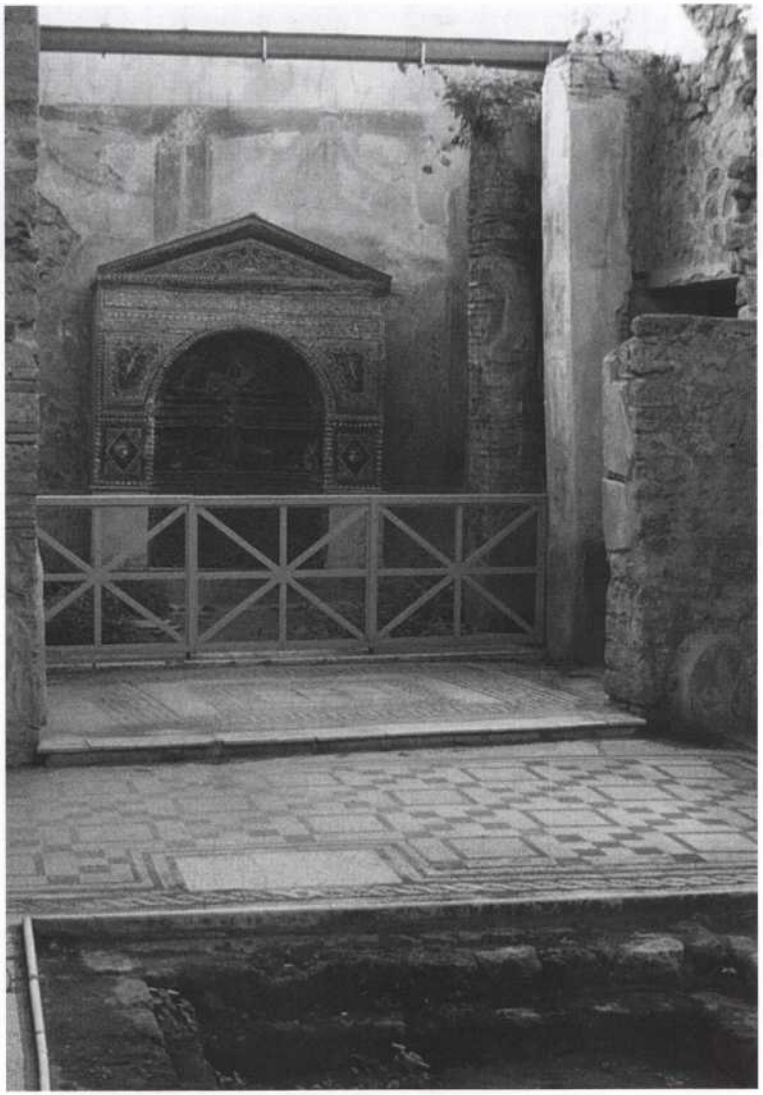


rapid degradation. Creating a larger space for culture on the Web would not only be important to the present scientific and public communities, but also help to secure cultural heritage against the threats of war and natural catastrophes for future generations. The Web could preserve digital representations of cultural objects for the memory of mankind and also serve to easily identify originals which have been lost because of theft or plundering, thus helping to minimize their potential on the antiquities market. Openly accessible catalogues of the cultural heritage of mankind kept in museums and other collections would therefore constitute a direct and powerful support of the UNESCO preservation policy whose urgency and viability the recent events in Iraq have made all too clear.

At present, however, the bulk of information which forms the core of cultural heritage, the great works of literature and art, as well the treasures of scientific, scholarly, and philosophical writings going back to the dawn of our civilization, are largely excluded from the information system already constituting the backbone of an evermore knowledge-based world. And the little culture that is included in the World Wide Web due to the efforts of a few pioneers is almost drowned by the tides of information garbage.

It is precisely the few shining examples of culture on the Web that make evident the potentials of the bulk of information constituting our cultural memory, which is still not represented within the new medium. Among these potentials is the chance of overcoming the fragmentation of cultural heritage by traditional institutions and disciplines. This fragmentation process has been determined, to a large extent, by preservation concerns (it is easier to preserve paintings by gathering them together in the same building, and the same holds true for the preservation of books, archival records, drawings, natural history objects, video tapes, sound tracks and so on). But, once the information has been processed and transformed into a digital entity, it must no longer follow the fate of its physical support. There is thus no reason to store it according to the same systems used to preserve the object it emulates. On the Web, there are no buildings or walls, and we are not obliged to reproduce distinctions based on the topologies of material objects or on the various nature of their physical shells. It is possible – and indeed necessary – to reorganize digital records into new cognitive architecture, where the strict constraints of the physical world no longer apply.

What is needed is a vision exploiting the new technological possibilities for the creation of a public culture of science, a vision that includes the humanities and thus keeps alive the roots of our

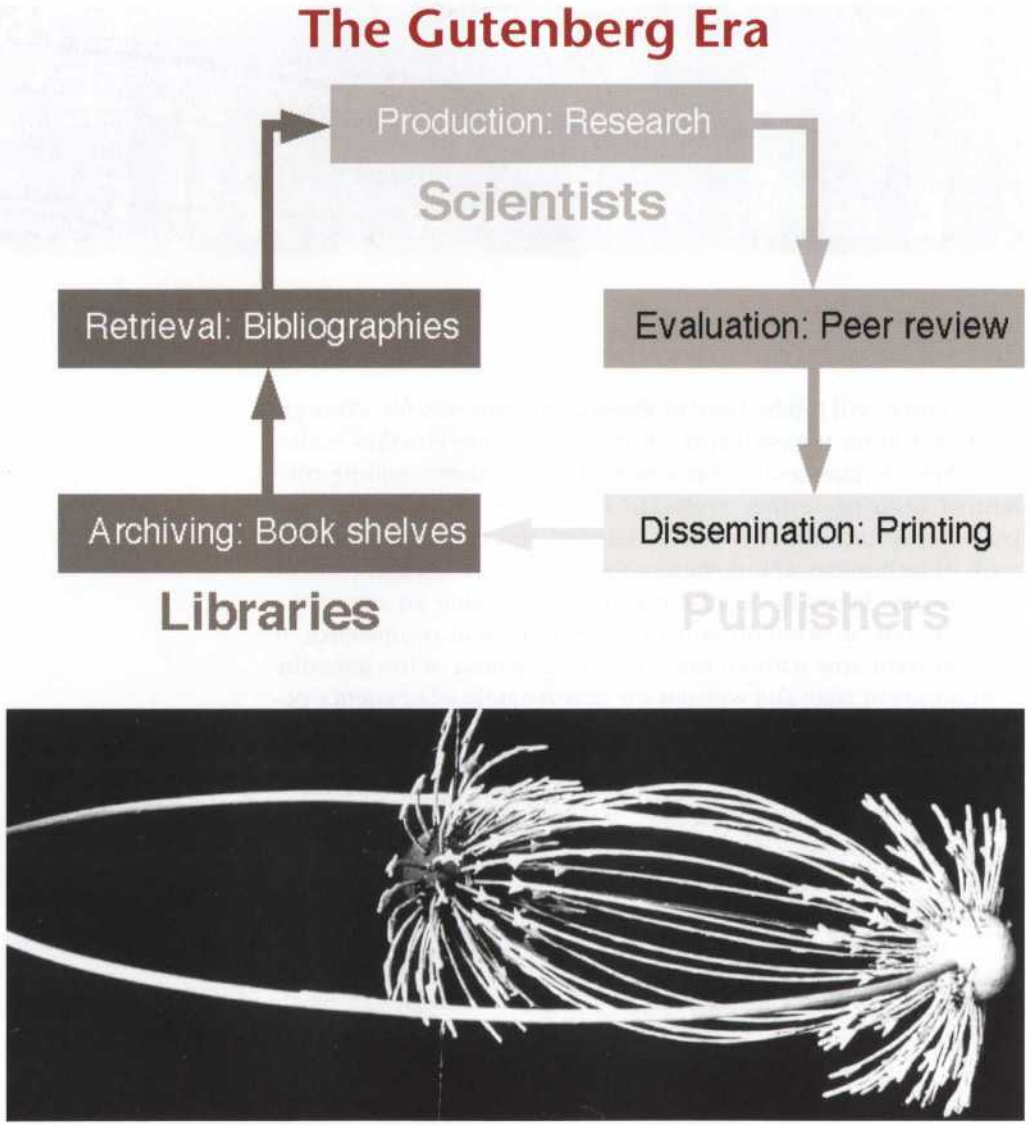


techno-scientific world in our cultural history. Such a vision must address a double challenge presenting itself to cultural heritage in the age of the Internet, a quantitative and a qualitative one: the need to make a substantial amount of the sources constituting the cultural memory of mankind electronically available, and the need to create an adequate intellectual, technological, and social infrastructure rendering this cultural memory accessible as a resource for addressing the questions of today, be they scholarly or from an orientation-seeking public.

The deficit in the extent to which cultural information is available on the net is indeed accompanied by another deficit with perhaps even more pernicious consequences: the underdevelopment of cultural techniques adequate to the new information technologies. In the fields of language technology, image analysis, and the implementation of mathematics on the net, that is, in fields of high economic and technological impact, bottlenecks become visible that are related to the negligence of an adequate transfer of the traditional cultural techniques of writing, depicting, and calculating to the new medium. The problems of language technology, for instance, have long been considered merely an engineering challenge and not a field to which the humanities can bring their century-long expertise in the linguistic representation of meaning.

Meaning is, after all, not only in the text but also in the cultural context so strikingly absent from the new medium. In short, the lack of implementation in the new medium of the cultural information and techniques which are the domain of the humanities represents a major stumbling block to what might otherwise become a second Internet revolution.

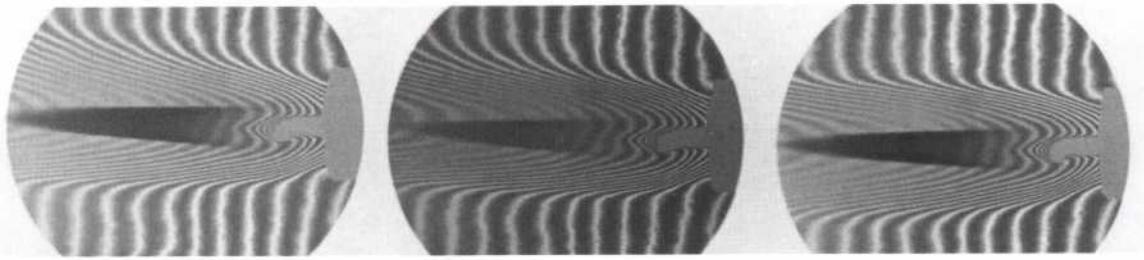
This revolution will, however, not take place automatically, merely as a consequence of technical developments, but it requires the creation of new, content-laden information structures which can only result from an effort to overcome the present marriage of ignorance between the scholarship and technology. What is needed is not just hardware and software, but a grid-like, open infrastructure supporting the accumulation and extraction of meaning from information distributed over the Web. An infrastructure capable of responding to these challenges will have to support preservation and free access to the cultural heritage of mankind against the ruthless and short-sighted pragmatism of technical and economic progress. It will have to provide public facilities for web-supported training and education. It will have to encourage the free exchange of information and arguments across currently existing social, political, and religious boundaries. It will have to guarantee a lasting memory of mankind comprising a representation of human history in a new form. And it will finally have to develop mechanisms for self-organisation and for the evaluation of the information it makes available. These requirements cannot be fulfilled by merely adding content or utilizing further evolution of existing technologies. The goals necessitate a massive, well-reflected and carefully concerted push of technological innovation, social organization and content enrichment. Precisely because the creation of such an infrastructure depends on scholarly as well as technological competence, it will not come true without the active participation of the scientific community at large and without the massive support of science policy.



1.2 Challenges of the Information Revolution for Science

The new electronic media of information production and dissemination are in the process of dramatically changing the conditions under which scientific information circulates. They will affect the infrastructure of science no less profoundly than the invention of printing. In the Gutenberg era of printed information, the responsibilities for the main parts of the flow of scientific information are clear: Research results are produced by scientists. They are disseminated by publishers and archived by libraries. Information is filtered by a process of evaluation performed by scientists (peers) and organized by publishers. Only that which survives this filtering process is being disseminated. Information is retrieved by scientists using bibliographical tools within an infrastructure offered by libraries. The system is well-established and has been impressively stable. It is now, however, endangered by technological changes with radical consequences. Even within the system of printed information these technological changes are felt by the increasing prices charged by publishers for dissemination, which scientific organizations are no longer able to cover.

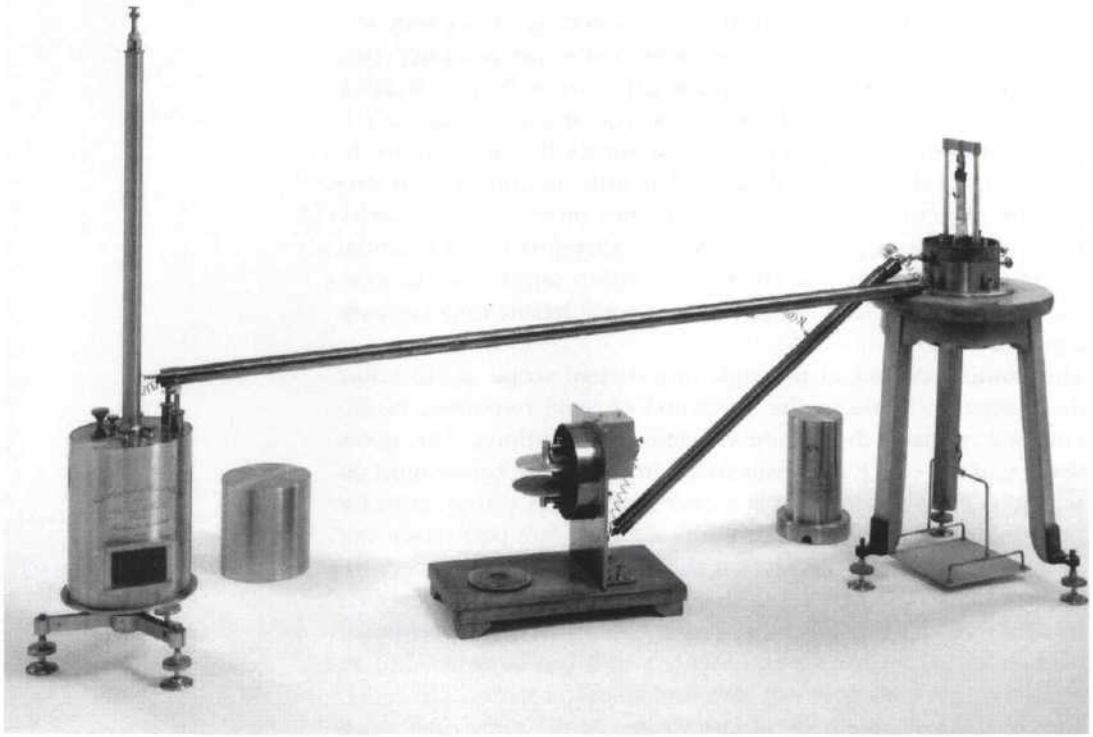
The information revolution has radically changed the technical and economic basis for maintaining the scientific information flow. Research results are being produced and can be immediately disseminated in electronic form. Dissemination is no longer a cost-intensive component. It can in principle be handled by scientists



without the services of the publishers. This is demonstrated, for instance, by electronic research archives for some areas of the natural sciences, which spread research results without any costs for the users once an institution is hooked up to the net. Furthermore, in the electronic medium evaluation follows and does not precede dissemination. This is the outcome of the self-reflecting capacities of a universal representation of knowledge. If this vision of the Internet era should lead to an acceptable model for the flow of scientific information, one has to make significant investments into the solution of two major open problems. The first is that of archiving, ensuring the long-term availability of electronic information. The second is that of an adequate access and retrieval infrastructure.

The forms of scientific representation will radically change as a consequence of the information revolution. There is no reason why future scientific representations should take the forms of journals or books, forms that are largely determined by the print medium and the respective agents. Instead they may take any form suitable as a contribution to a global scientific information network and its structure. Even now we are familiar with a variety of possible forms of representation, ranging from entries in data-bases, via digital archives and collections of links, to interactive research environments. There is, in particular, no longer any reason to preclude access to the information hinterland (observational and experimental data, software tools, historical sources), presently only serving as a logistic background for published research results. This will help ensure the reliability of scientific information, to broaden the scope of available resources, and to avoid the duplication of efforts. The electronic medium offers scientists the opportunity to reach, practically without delay and at little additional costs, the primary objective of their work, an impact on the body of knowledge of their scientific community. It is therefore in their natural interest to freely broadcast their information rather than to force their fellow-researchers to buy it with a considerable time-lag from a publisher.

The immediacy and in principle unrestricted scope of electronic dissemination increases the likelihood of rapid responses, distinguishing valuable from non-valuable contributions. The quick settling of the cold fusion issue in the Internet even before most discussions appeared in print is a case in point. Of course, even for print information it is, at least in the long run, not peer review but usage that eventually decides on the quality of a scientific contribution. Under the new conditions, the same process can exercise its selective effect much more rapidly. In contrast to traditional peer-reviewing, open peer-commentary as it can be realized in an electronic network does not lose valuable information but rather adds to the available body of knowledge. At the same time, open



peer reviewing allows for more differentiated judgements rather than just for an "in/out" decision about publication which does not distinguish quality differences between materials that have survived the selection process. The new medium could thus facilitate and improve the quality of the selection process. An appropriate infrastructure for transforming quality judgement into a navigational aide in the ocean of information is, however, at present still not available.

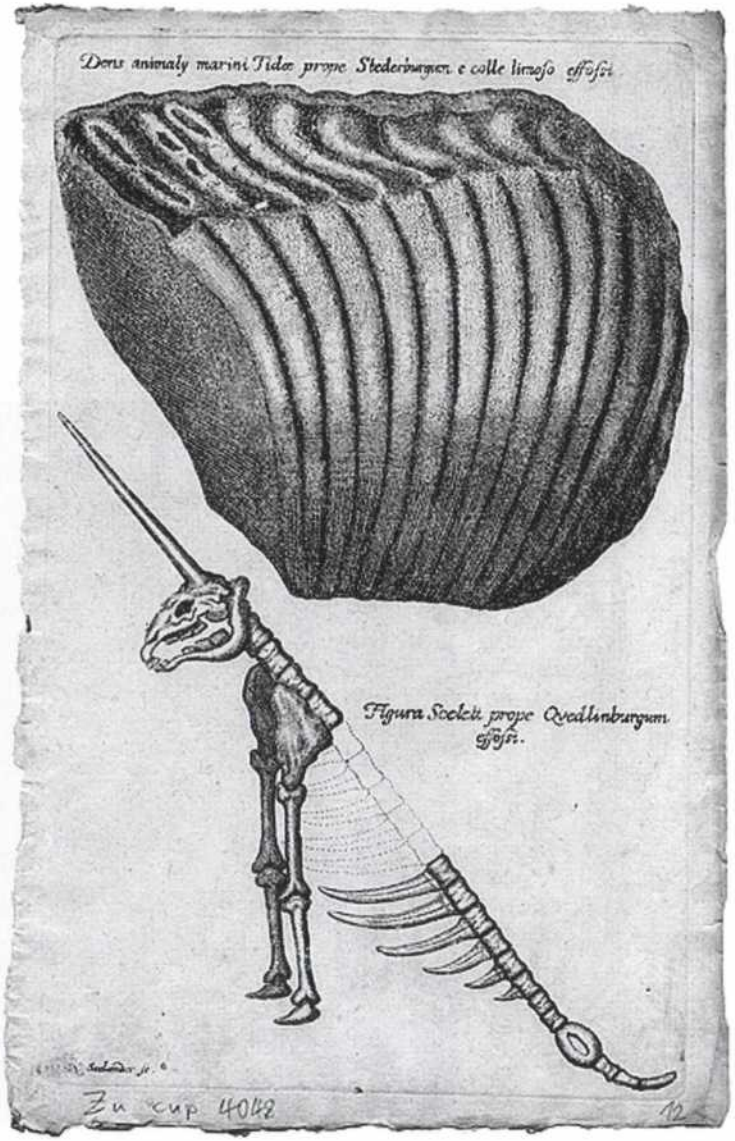
The question of how to ensure the longevity of electronically represented scientific information is, of course, not only a technical but also an institutional problem. No single type of institution is presently equipped to offer a complete solution. Most probably, only a kind of "New Deal" among the agents in the scientific information flow can provide the basis for a solution. In any case, this problem represents a major challenge that has to be addressed by research organization not only at an administrative but also at a political level.

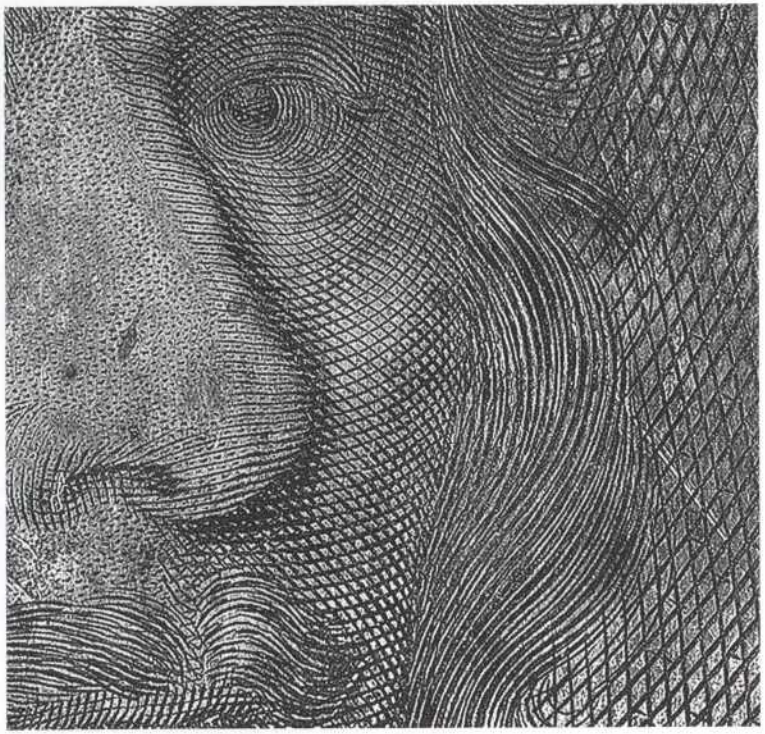
Since the costs for information dissemination have been dramatically reduced, the publishers will in the long-run risk losing their main source of revenue unless they offer new services, adding value to the scientific information. The new role of libraries in the Internet era is also open. They have to find their place in the new distribution of labour.

1.3 The Insufficiency of Existing Solutions

How can the pernicious situation described above be changed? Different scenarios are conceivable: The "scout solution" is based on the assumption that the transfer of cultural heritage to the new medium can be achieved by pilot ventures only. The other common scenario is the "big player" solution. It essentially assumes that the dominating forces of the economic or academic market will sooner or later take care of bringing cultural heritage to the net.

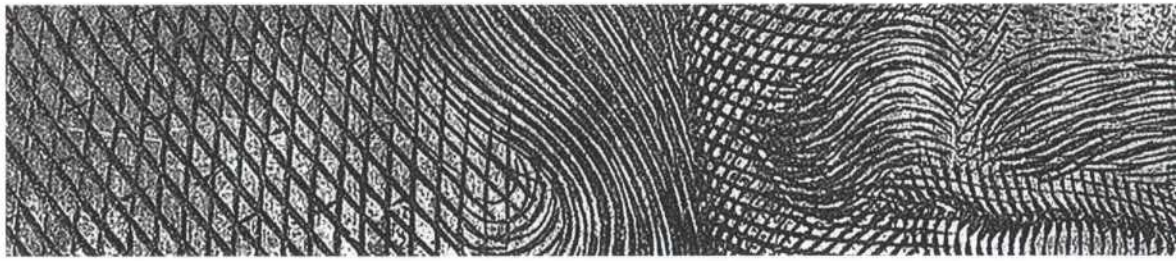
The big player solution is most familiar from present debates on electronic journals. While the few publishers who hold a near monopoly in certain areas of scientific publishing are indeed offering more and more material on the Internet, their approach has been rightly characterized as a "Faustian deal" in which a fatal price has ultimately to be paid by the scholarly community. In fact, although electronic dissemination is considerably cheaper than print dissemination, journal prices — in general still coupled to those of print subscriptions — not only continue to increase but, what is worse, the revenue accumulated by the publishers is in general not reinvested in a future-bound infrastructure for scientific information on the Web. On the contrary, the great challenges for such an infrastructure, for example, the archiving problem or the problem of an integrated retrieval environment remain, for the time being, largely unsolved — menacing the longevity and inter-





operability of scientific and scholarly information in the electronic medium. This is the fatal price of the Faustian deal. There will be no escape from it as long as the scholarly community has to repurchase from the big players the information it produced in the first place, at the same time being left responsible for its infrastructure on the Web.

The situation of the digital availability of the primary sources of cultural memory is even more problematic. While sceptics are still debating the compatibility between culture and the Web, the big players have long since begun to secure exclusive rights on the reproduction of cultural artefacts and even to purchase important documents and collections with the intention of commercialising their digital images. From the codices of Leonardo da Vinci to the photographs of Ansel Adams, every piece of cultural heritage is a potential asset in this new market. In the hope of spectacular gains, new firms have been founded, claims staked out, and portals opened up with a "gold rush" mentality. And indeed, looking back at its first phase, one already recognizes the typical ruins documenting the transiency of every gold rush: portals promising to become gateways to unimaginable cultural treasures which actually lead nowhere; key documents of European history are, on the other hand, confined to CD ROMs which are condemned to gather dust until they become outdated with the next generation of soft



or hardware. Meanwhile they are banished from the World Wide Web, which enlivens and enhances every significant piece of information exposed to it by its self-organizing connectivity, at least as long as this connectivity is not smothered by passwords or pay-per-view access. It has become particularly evident that the big players have failed, in spite of their eagerness to control large domains of cultural heritage, to create an infrastructure that guarantees a steady and reliable flow of this heritage from the old medium into the new. On the contrary, they have contributed to an increasing inaccessibility of cultural heritage — not only because of the restrictive copyright laws they seek to impose but also because sources are now often held back by museums, archives, and libraries in the dim hope of future commercialisation. This hope can, however, hardly be sustained by a practice that amounts to a ruinous exploitation of limited resources rather than representing a concerted effort to augment them.

The scout solution, on the other hand, is based on the assumption that the transfer of cultural heritage to the new medium can be achieved by pilot ventures, perhaps in combination with an establishment of standards for production and dissemination. In contrast to the big player solution, it amounts to the realization that bringing culture to the Internet actually means settling a new continent rather than just exploiting its resources in a gold rush. But it also amounts to the assumption that this can be done by merely sending out a few scouts to survey the new territory and set up a model farm here and there. However, this approach should not be criticized too brashly. As a matter of fact, almost everything presently available in terms of digital libraries demonstrating the potential of the new media for cultural heritage is due to the breakthroughs achieved by this strategy. But it must be legitimate to ask whether this strategy is adequate to meet the principal challenge of the future: the creation of a self-sustaining representation of culture and science in the new medium.

Looking back at the successes and failures of the projects funded by national agencies as well as by the European community, one finds indeed that so many of the feasibility studies, pilot projects, test beds, and proofs of concept, however impressive they are if taken by themselves, have actually failed to launch such a self-sustaining dynamics. The dead links, blind alleys, and empty databases characterizing some of the most ambitious homepages of such projects signal that they did not succeed in making a difference for the scientific community at large, let alone for the role of cultural and scientific memory in an Internet society. Such projects are rather like chip factories in the jungle, incapable of leading off productive development because even the most basic infrastructure is lacking.

Admittedly, the humanities, responsible for preserving, exploring, and keeping cultural heritage alive, is a difficult environment for technical innovations. Scholars in the humanities have hardly even begun to realize that the new information technologies not only confront them with a competence problem that is unparalleled in the natural sciences but also that they are faced with entirely new possibilities to overcome the deeply entrenched boundaries of narrow specialization. The assumption that the humanities can be catapulted into the Internet age by enticing them with exemplary pilot projects reminds a historian of the overlay astute attempts of Jesuits in the 17th century to convert the Chinese mandarins to Christianity: they offered in fact a few extraordinarily beautiful clocks as a gift to the Chinese emperor in the futile hope that he would ask for more European technology and religion once the donated clocks needed rewinding.