MAX-PLANCK-INSTITUT FÜR WISSENSCHAFTSGESCHICHTE
Max Planck Institute for the History of Science

PREPRINT 151 (2000)

Jürgen Renn

Challenges of the Information Revolution for the Max Planck Society

CHALLENGES OF THE INFORMATION REVOLUTION FOR THE MAX PLANCK SOCIETY[1]

Jürgen Renn

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. Let us therefore take a look at this infrastructure in the era which was dominated by printed information, the ”Gutenberg Era” (see fig. 1).

THE GUTENBERG ERA

In the 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 what survives this filtering 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 whose radical consequences I will address next. Even within the system of printed information they are felt by increasing prize-charges requested by publishers for dissemination charges which scientific organizations are no longer able to cover.

---

[1] Lecture given at the conference of the Max Planck Society „The Transformation of Science, Research between Printed Information and the Challenges of Electronic Networks,“ which was held at Schloß Elmau, May 31 - June 2, 1999.

THE GINSPARG ERA

Let me confront this brief description of the status quo with a vision of what will be the outcome of the technological revolution in the field of information processing and communication which is going on at the moment. I call this vision ”the Ginsparg Era“ (see fig. 2). This 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 Paul Ginsparg’s Los Alamos e-print-server (http://xxx.lanl.gov), an electronic research archive for physics, which spreads 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. It is the outcome of the self-reflecting capacities of a universal representation of knowledge.

If this vision of the Ginsparg 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.

Let me stress from the very beginning the main intention of what I am going to say: We are at a bifurcation point between directions into the future. We either have to invest into the realization of a new information management system for the sciences. Or else we have to strongly fight against the ongoing dissolution of the traditional model by the infiltration of electronic means into the information circulation system, a process which is going on, as will become clear in the following, for merely economic reasons. Furthermore, I am personally convinced that, if the Max Planck Society does not take a decision on this question, the society simply has to accept whatever direction the current revolution will take.

CHALLENGES OF THE INFORMATION REVOLUTION

Let me briefly comment on the main challenges of the information revolution for a scientific organization and its partners in the circulation of scientific information.

INFORMATION WILL BECOME A PATCHWORK-LIKE STRUCTURE.

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, textual sources), presently only serving as a logistic background for published research results. This will contribute to ensure the reliability of scientific information, to broaden the scope of available resources, and to avoid the duplication of efforts.

INFORMATION WILL BE BROADCASTED.

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.

QUALITY CONTROL WILL BE REGULATED BY USAGE.

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 – or if you prefer ”history” – 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 loose 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 missing.

ARCHIVING IS AN UNSOLVED PROBLEM!

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 such as the Max Planck Society not only on an administrative but also on a political level.

PUBLISHERS AND LIBRARIES WILL HAVE TO PROVIDE NEW SERVICES!

Since the costs for information dissemination have been dramatically reduced, the publishers will risk in the long-run to loose their main source of revenue unless they offer new services, adding value to the scientific information. The new role of libraries in the Ginsparg era is also open. They have to find their place in the new distribution of labor. Clearly, the separation between e-worlds and print-worlds, still common in scientific institutions such as the MPG, is challenged by this new distribution of labor.

Present Structure of Information Management in the MPG

After having reviewed some of the challenges of the information revolution, let me briefly remind you of the current situation of information management in the MPG. Due to the federal structure of the MPG, one has to distinguish between local and central aspects. On a local level, there are, in addition to the services of the libraries and computer support units of about 80 institutes, various advanced information management projects, such as the development of a research project server at the Fritz Haber Institute or the EUDICO project on language corpora of the Max Planck Institute for Psycholinguistics. On a central level, this support is complemented by two information retrieval units, the IT department of the Administrative Headquarters, and the Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG).

The central advisory board on IT of the president of the Max Planck Society is the BAR which, in spite of its anachronistic name ”Beratender Ausschuss für Rechenanlagen,” has shaped the IT policy of the society in recent decennia. It was also the BAR which initiated the first steps towards defining a new information management policy for the Max Planck Society. In particular, an information task force has been founded to realize an intermediate solution for the e-journals problem. The present structure of information management has offered in the past optimal conditions for research.

Limits for the Exploitation of New Technologies

The present structure includes, however, also serious drawbacks in the present crisis. In particula, an infrastructure necessary to generalize the developmental dynamics is lacking. In some fields we observe a considerable developmental dynamics: In the life-sciences, in physics, in the geological sciences, and in astronomy, for instance, large factual data-bases of research results are available world-wide for a global communication on their interpretation. In other fields, and in particular in the humanities, on the other hand, the existing potential of the new media is not exploited with a comparative dynamics. What is missing is an infrastructure that allows to generalize this developmental dynamics. As a consequence, it often looks as if the ”wheel is re-invented” over and over again.

Similarly, a response to the question of determining suitable dissemination strategies for research results requires an appropriate infrastructure which at present does not exist.

Last but not least, as a major research organization, the Max Planck Society has to make an important contribution to shaping the future infrastructure of science and to protect the access to scientific information in the new division of labor. But it needs an appropriate infrastructure in order to raise its voice in the political discussion on the protection of scientific information. Such an infrastructure at present only exists in a rudimentary form.

WHAT IS *NOT* BEING DONE

In order to illustrate why I think that the Max Planck Society does at present not adequately take up the challenges of the new technologies, let me mention some examples:

Software-instruments for the management of research information as they are needed by the institutes such as project-servers, digital archives, tool-boxes, data-mining tools and the like are not being developed other than in exceptional cases. One of the reasons is that neither the single institutes nor central institutions have, in general, sufficient resources for such tasks. As a result of a survey of the BAR, we now know that more than half of the institutes have not even a single person responsible for software development, and a quarter of the institutes have only a single person available.

Solutions for standard problems such as the structuring of electronic texts for research purposes or the documentation of research materials, even if they exist somewhere in the society, are not systematically made available to the institutes. There is no infrastructure offering advice concerning their implementation.

In view of the new ways in which research results can be made available, the development of novel dissemination and evaluation models becomes urgent but at present there is no infrastructure for supporting such efforts. In particular, new institutional structures are needed for dealing with the problems of electronic journals, databases, etc., such as those concerning policy, provision, and support.

WHAT TO DO?

So, what are the main questions about which the Max Planck Society has to decide now:

- Can the society just wait and see?
- Is it enough that the Max Planck Society provides its institutes with the best commercially available information services?
- Which support do the institutes need in order to play the part they have to play in taking up the present challenges?
- And if support is going to be provided, where shall the funds come from?

In order to answer this last question let us take a bird‘s-eye view on the financial aspects of the circulation of scientific information.

THE GUTENBERG ERA – COSTS

I begin with the familiar situation in the Gutenberg Era (see fig. 3). The total annual budget of the Max Planck Society which represents the total costs of the scientific information produced by the society is about 1, 3 Billion $. The society spends annually about 10 Million $ for journals which still represent the primary access to current information generated by the entire system of scientific institutions. That is, in the Max Planck Society, the total cost of access to current information is in the order of magnitude of one percent of that budget, a sizeable fraction.

THE GINSPARG ERA – COSTS

Of course, it is difficult to say anything definite about costs in the Ginsparg Era (see fig. 4). But one thing is certain: after a successful revolution, the fraction of costs for the access to information compared to those for its production will be drastically reduced. Under these conditions, each scientific institution will take care itself of the electronic dissemination of the information it produces. Printed journals will have vanished. On the basis of a conservative estimate, the dissemination costs additional to those for the already available equipment and services will probably be well below 1 Million $, that is, less than a tenth of the present expenditure.

But the total costs for implementing such a solution will have to include those for solving the two major structural problems mentioned above, that of archiving and that of retrieval. As long as no concrete solutions for these problems have been envisaged for the Max Planck Society, nobody will be able to give reasonable estimates about the possible costs. It should be clear that if one wants to go in the direction of the Ginsparg vision, hand-in-hand with the change of the dissemination system, the funds that are presently spent for the provision of scientific information have to be redirected into new technologies for archiving and retrieval.

So much for future perspectives. Let us leave the grand visions for a moment and turn to what is going on at the moment. The Max Planck Society has taken first measures to react to the e-journals-offer of the publishers. The information task force has, together with the IT department of the Administrative Headquarters, realized the pilot phase of e-journals, covering 1600 journals for an approximate annual amount of ca. 330 000 $. This experience has allowed us to make first predictions concerning what the global introduction of e-journals in the Max Planck Society would cost under the condition of the current test phase.

THE E-JOURNALS DEAL – COSTS

In fact, also the publishers have realized that the costs for electronic dissemination is much lower than for printed dissemination. Although the estimated figure of 3 Million $ covering all journals presently subscribed to by the Max Planck Society is higher than the estimated dissemination costs for scientific information in the Ginsparg vision, it is still considerably less than the 10 Million $ which the society presently spends on printed journals. However, this offer of the publishers is valid only under a certain condition, on which it seems, from our experience with tough negotiations, that there will be no way of altering the position of the publishers. The publishers are willing to offer their e-journals for the present prize only if the Max Planck Society maintains its subscriptions to the printed versions to essentially the same extent as presently. This is what I would like to call ”the e-journals deal” (see fig. 5). To be completely honest, this deal meets also to a considerable extent the expectations of the scientists of the Max Planck Society. As we know from a poll of the Information Task Force, whereas on the one hand the offer of electronic versions is highly appreciated, there is, on the other hand a considerable resistance to cancel the subscriptions to the parallel printed versions.

THE FAUSTIAN DEAL – COSTS

Let me finally come to what the editor of *Behavioral and Brain Sciences*, Stevan Harnad, has called the "Faustian bargain" (see fig. 6). Given the situation as it is, what would happen if the Max Planck Society would adopt a policy to cancel all subscriptions to printed versions as soon as an electronic version becomes available? In this case the total cost for access to content would be reduced. In fact, this would, at least at present, be the cheapest solution for the Max Planck Society. From the negotiations it has become clear that the publishers would reduce the total costs after substitution of the printed by the electronic versions of the journals by an order of magnitude of at most 20%. As a rule the publishers either insist on a political prize to maintain the total amount of their revenue or reduce it by about 10%. There are many reasons given for this policy which, however, I am not willing to discuss here.

This is in fact a Faustian deal because the prize has to be paid at the end. Firstly, the costs for electronic journals will not cover the archiving of their contents. Secondly, the Max Planck Society would leave to the publishers any decision on the policy of the dissemination of scientific information.

MODEL SOLUTIONS

What is the gist of this account for possible strategies the Max Planck Society could adopt?

THE TITANIC SOLUTION: ALL RESOURCES INTO SERVICE!

The Max Planck Society could adopt the Titanic Solution and invest its resources for services obtained under the conditions of the Faustian deal, hoping not to hit the iceberg in the end. There may even be a good chance.

THE PATHFINDER SOLUTION: LET THE INSTITUTES FIND THEIR WAYS!

The Max Planck Society could adopt the Pathfinder Solution and let the institutes find their individual ways to deal with the problem. They might decide individually how to disseminate the information they produce and how they get access to the information provided by other institutions.

They might continue to subscribe to an ever smaller number of print journals and, in addition, rely on the ever increasing electronic information resources freely available. In any case, it should be clear that, in this case, the institutes need more support, partly for development, partly for paying the increasing journal prizes.

The Laager Solution: centralize all forces!

The Max Planck Society could adopt the Laager Solution and centralize all forces. The society might think of concentrating the existing central services and other resources for information management in one place. This might lead to an increase of efficiency by improved coordination. The comparison with other scientific organizations would suggest the creation of a central unit of considerable size.

The Midwife Solution: combine central services and support with local developments of innovative solutions!

Let me end with the strategy that I personally find the most promising. The Max Planck Society could adopt the Midwife Solution and combine enhanced central services and support with local developments of innovative solutions. The society might think of temporarily creating a center where a component supplementing existing services and a flexible task force with developmental tasks are combined. The task force could flexibly support local information management initiatives and then transform these local experiences into the basis for the development of generalized solutions. In close cooperation with the Administrative Headquarters, the service component could take care of the organization and policy necessities of information provision in the present transitional phase. At the same time, in cooperation with the development component, it could create models for long-term solutions of information provision to the society and rapidly implement them where possible. An example might be the creation of a Max Planck Server for electronic publications.

I mentioned in the beginning that we are at a bifurcation point between directions into the future, either fighting to preserve the traditional system or invest into the realization of a new information management system for the sciences. It is clear that all options we have considered here fall under the latter category. Fighting against progress or just bowing to economic restrictions is not compatible with my understanding of the Max Planck Society. In any case, it seems clear that just sit and wait will not be adequate to meet the challenges outlined above.

Fig. 1

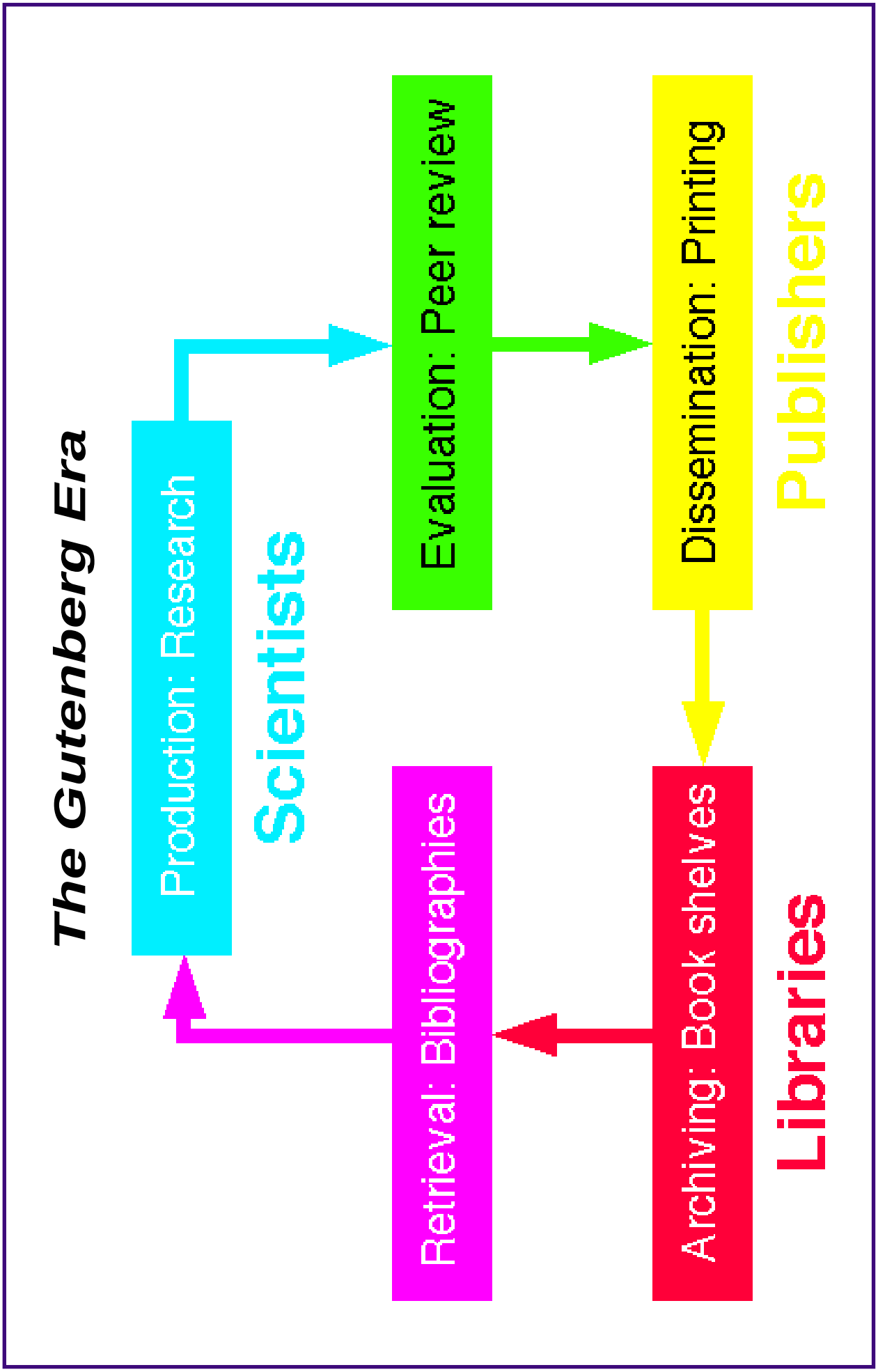

The Gutenberg Era
Production: Research
Scientists
Evaluation: Peer review
Dissemination: Printing
Publishers
Retrieval: Bibliographies
Archiving: Book shelves
Libraries

Fig. 2

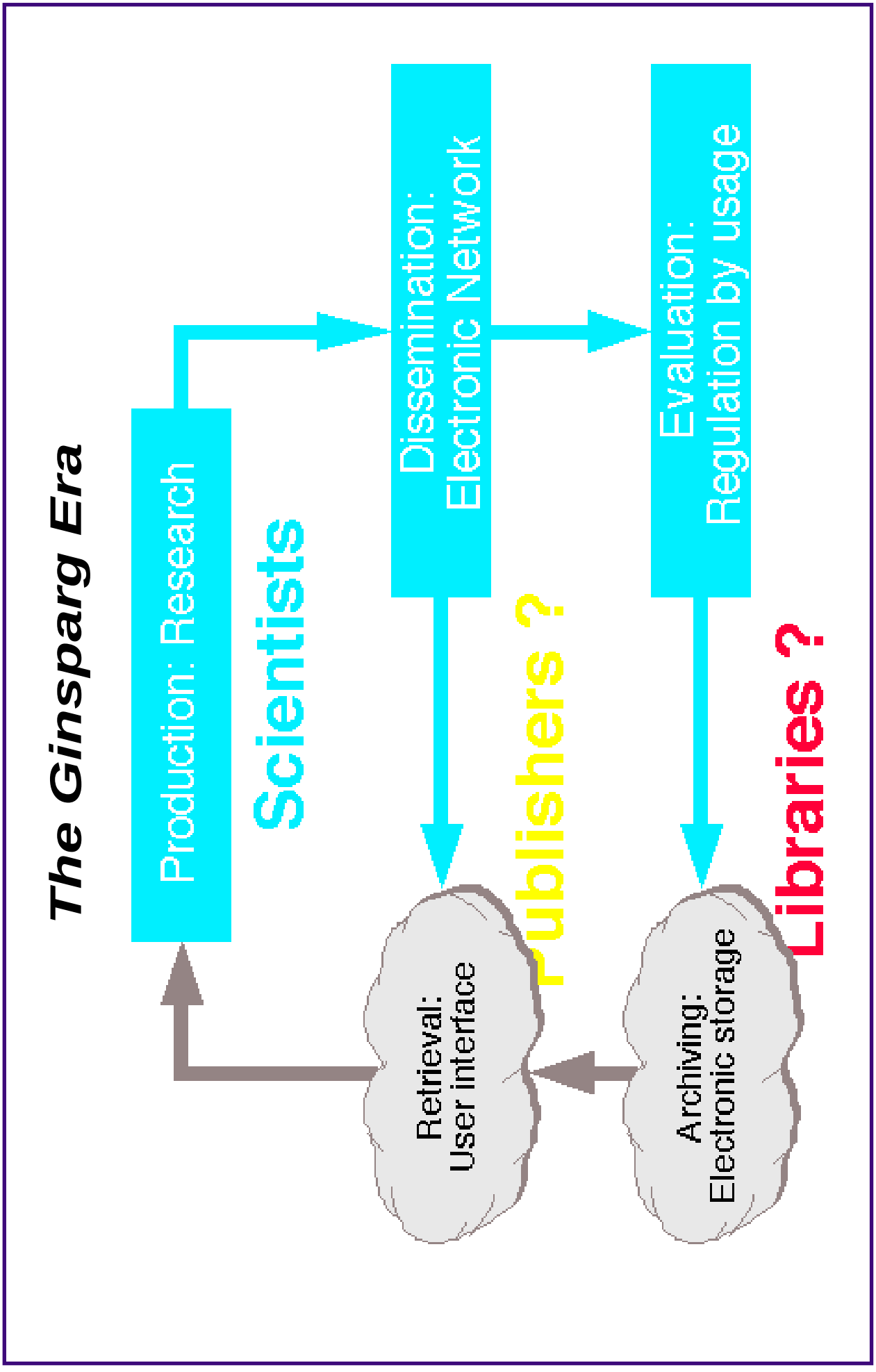

The Ginsparg Era
Production: Research
Scientists
Dissemination:
Electronic Network
Evaluation:
Regulation by usage
Publishers ?
Libraries ?
Retrieval:
User interface
Archiving:
Electronic storage

Fig. 3

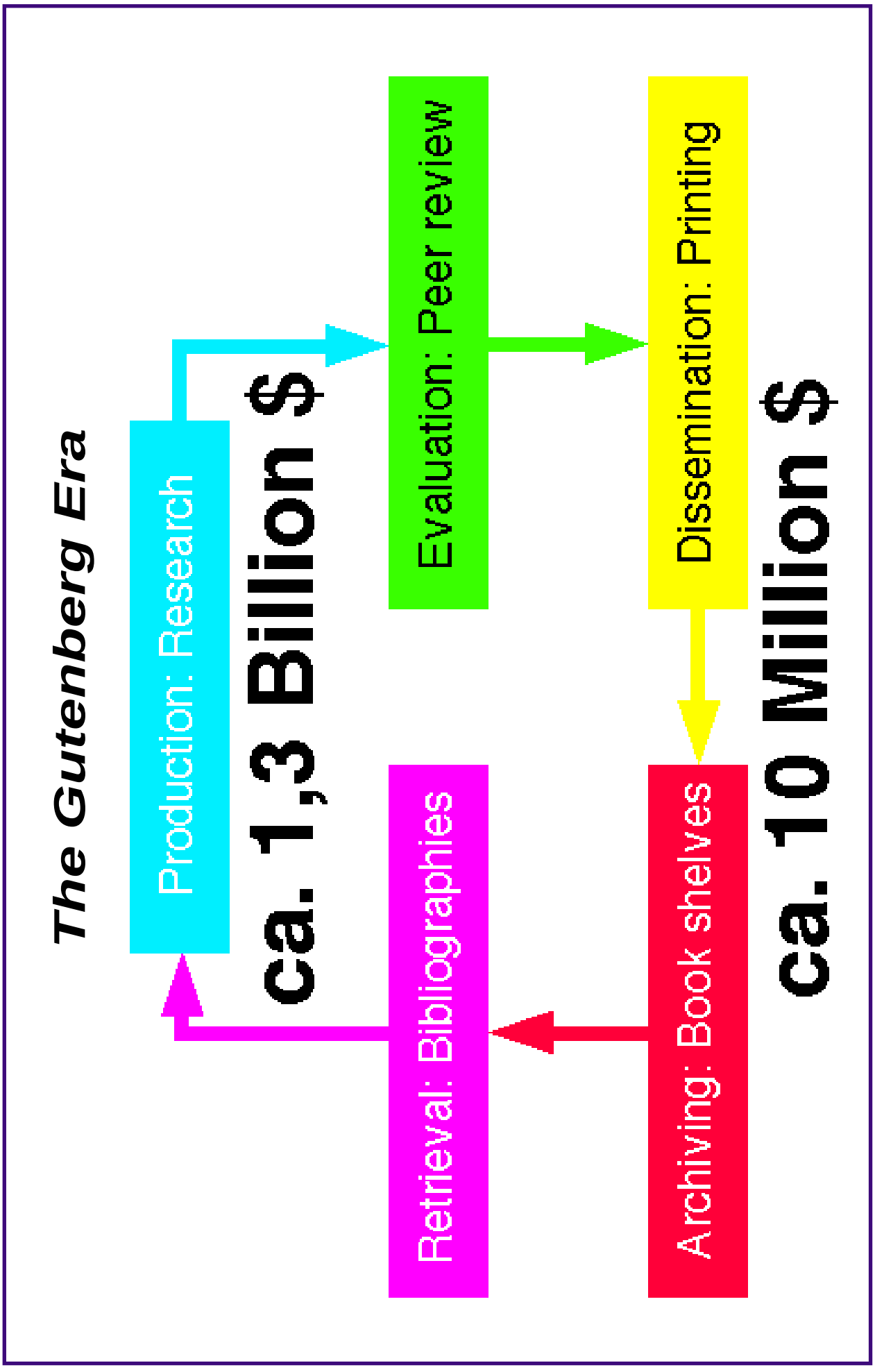

The Gutenberg Era
Production: Research
ca. 1,3 Billion $
Evaluation: Peer review
Dissemination: Printing
Retrieval: Bibliographies
Archiving: Book shelves
ca. 10 Million $

Fig. 4

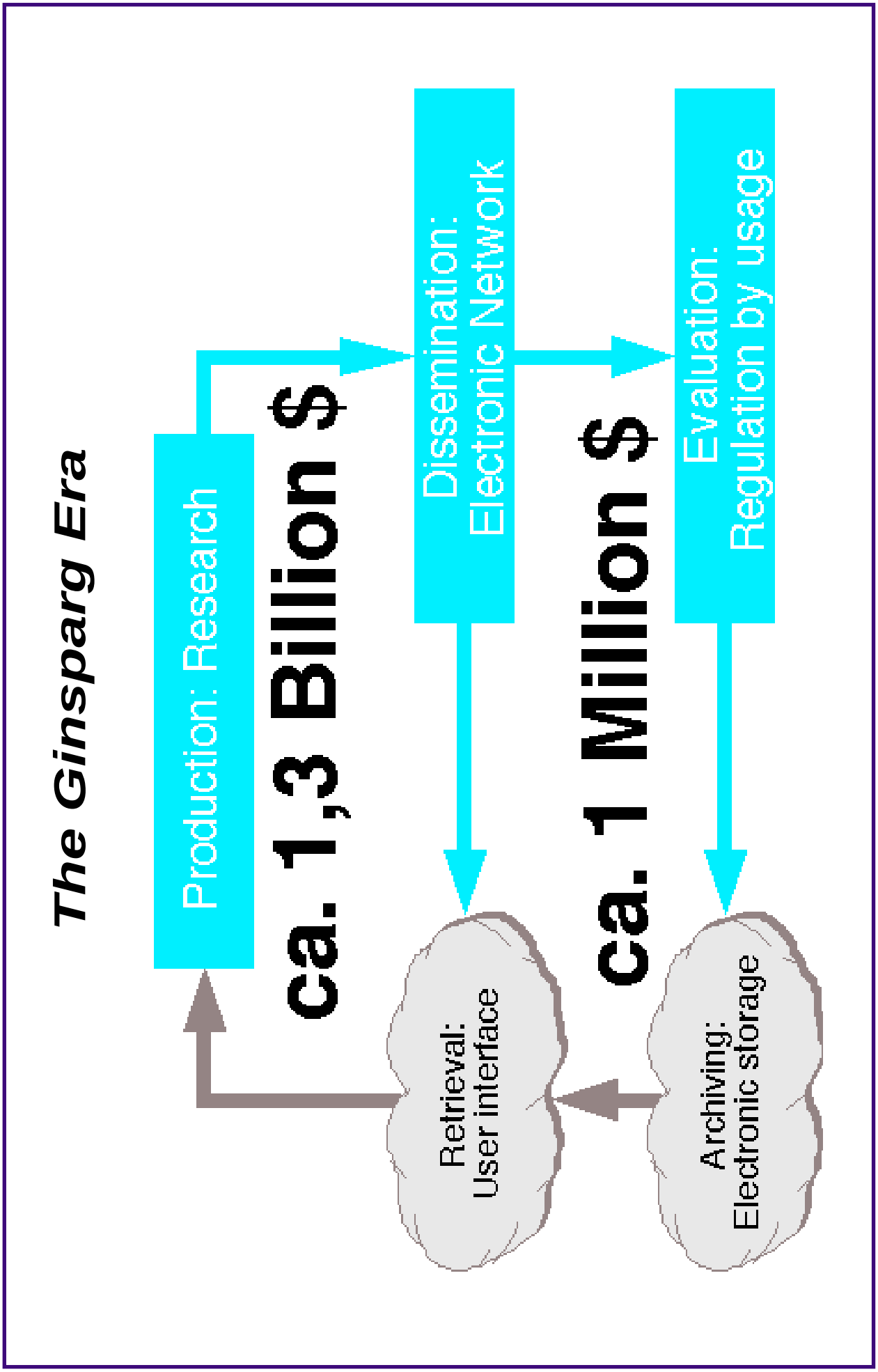

The Ginsparg Era
Production: Research
ca. 1,3 Billion $
Dissemination:
Electronic Network
ca. 1 Million $
Evaluation:
Regulation by usage
Retrieval:
User interface
Archiving:
Electronic storage

Fig. 5

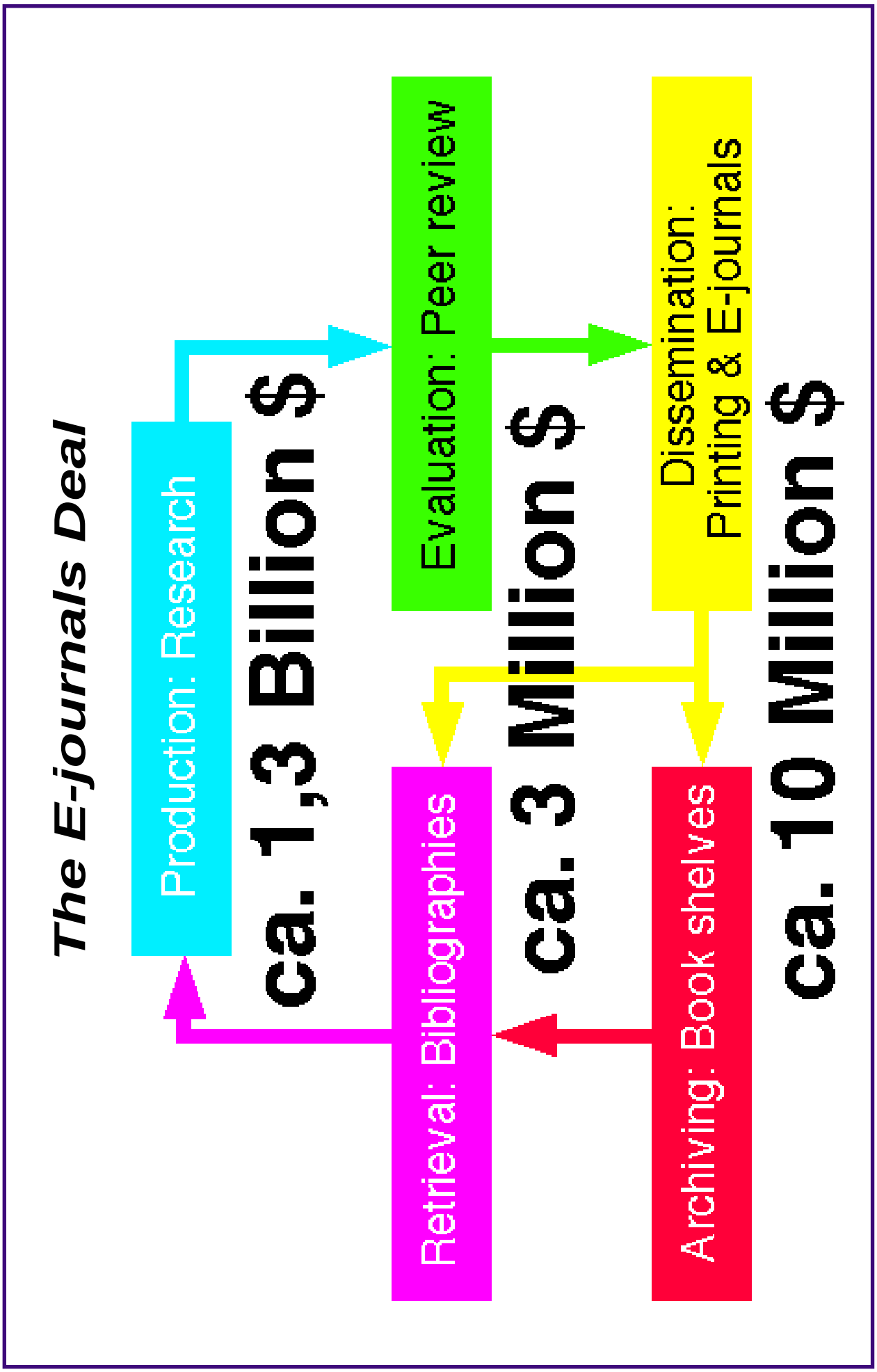
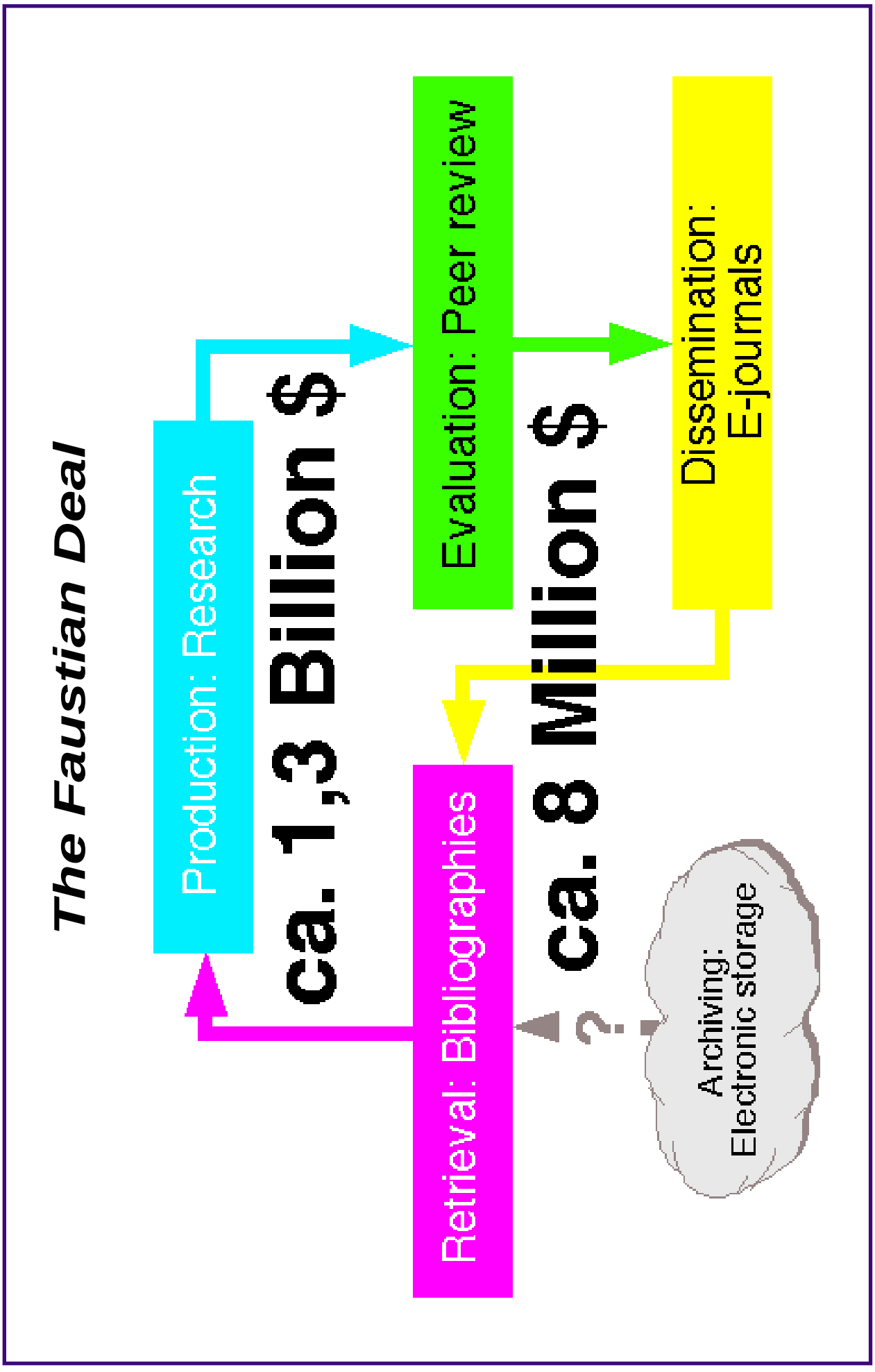

The E-journals Deal
Production: Research
ca. 1,3 Billion $
Evaluation: Peer review
ca. 3 Million $
Dissemination:
Printing & E-journals
ca. 10 Million $
Retrieval: Bibliographies
Archiving: Book shelves

Fig. 6


The Faustian Deal
Production: Research
ca. 1,3 Billion $
Evaluation: Peer review
Dissemination: E-journals
ca. 8 Million $
Retrieval: Bibliographies
?
Archiving: Electronic storage

APPENDIX

The following document served as the central discussion paper for the conference of the Max Planck Society „The Transformation of Science, Research between Printed Information and the Challenges of ElectronicNetworks," which was held at Schloß Elmau, May 31 - June 2, 1999.

A Center for Information Management of the Max Planck Society (CIM)

**This proposal is the result of discussions of the Steering Group of the Project Electronic Information Provision of the Max Planck Society, Ian Baldwin (MPI for Chemical Ecology),
Jürgen Renn (MPI for the History of Science),
Robert Schlögl (Fritz Haber Institute)
with Stephen Levinson und Peter Wittenburg (MPI for Psycholinguistics)**

The technological revolution in the field of information processing and communication can only be compared to the introduction of writing and to the invention of book-printing. It confronts science with two closely interrelated challenges: that of the reflection of this development and that of making use of it. In particular, disciplines dealing with the structures of social and cognitive systems are challenged not only to reflect the restructuring processes of science accompanying the information revolution, but also to shape it.

To date the humanities, but also the behavioral and social sciences have not been able to make use of the rapid dynamics of information technologies to a similar extent as is the case in the sciences. As the humanities compete with „junk information," the electronic representation of the cultural knowledge embodied in their disciplines is playing an ever decreasing role. Electronic archives, multimedia language corpora, or electronic translation interfaces - to name just a few examples - are developed and used occasionally, but they are far from forming an integrative part of the research routine in the humanities.

Even in the sciences the new possibilities of knowledge representation offered by the electronic media are up to now used only within the framework of a traditional research infrastructure. In some areas there are isolated innovations based on the new information technology, but its potential is not yet generally used for advancing the structure of the scientific working process. In the life-sciences, in physics, in the geological sciences, and in astronomy large data-bases of research results are

available world-wide for a global communication on their interpretation. Learned societies like the *Deutsche Physikalische Gesellschaft* begin to publish research reports in electronic form only. But since an infrastructure that allows to generalize this developmental dynamics is still missing, it often looks as if the “wheel is re-invented“.

Although the technical conditions for access to scientific information, its processing within the scientific research process, and its dissemination have changed radically, generally established solutions are still missing for such basic problems as the standardized structuring of electronically represented scientific information, the development of instruments for their processing according to the needs of researchers, establishing long-term solutions for archiving electronically stored information, determining suitable dissemination strategies for research results, and finally the optimal provision of scientific institutions with electronically represented information.

In view of this transitional situation, research organizations like the Max Planck Society have not only to redefine their information provision policy, but they have to make an important contribution to shaping the future infrastructure of science. They have to raise their voice in order to represent the specific needs of science in a process of transformation which is determined by a manifold of interests not necessarily congruent with those of science. Until now the Max Planck Society has been unenganged in initiatives and discussions of learned societies and other bodies on information management policies. Furthermore, it so far was no anchor for joint international funding initiatives. The wide scope of the disciplines represented within the Max Planck Society and the resulting variety of challenges in the use of electronically represented scientific information stand for a unique innovation potential which should be used for the achievement of this task. By coordinating and concentrating the expertise in the area of information management, the Max Planck Society can enhance its possibilities to successfully defend its interests with regard to external providers of scientific information such as commercial publishers (collective bargaining) and at the same time make optimal use of its internal resources.

A Center for Information Management of the Max Planck Society should consist of a component for developing long-term models for central services (e.g. the provision of the Society with electronic journals and data-bases), and a flexible task-force, supporting Max Planck Institutes in their local information management tasks and transforming these local experiences into the basis for the development of generalized solutions. The combination of these two components should guarantee that in a rapidly changing environment services are always up-to-date and that innovative solutions created by the task-force can be broadly implemented. Through a special funding within the framework of the „Research Initiative“ program of the Max Planck Society, institutes can - together with the Center - realize projects requiring advanced instruments for information management.

In close cooperation with research projects of Max Planck Institutes and also with existing outside iniitiatives, the Development Component of the Center aims at

- enlarging the access to research resources
- implementing existing international standards
- developing software-instruments for the management of research information when appropriate tools are not available

- making available solutions for standard problems and advising the institutes concerning their implementation
- participating in the development of standards for structuring complex research data
- participating in developing novel dissemination and evaluation models for research results
- participating in finding solutions for the long-term availability of electronically represented scientific information
- developing a Max Planck server for electronic publications

In close cooperation with the existing information management facilities of the Max Planck Society, the Service Component of the Center aims at

- providing the Max Planck Society with electronic journals, databases, and other electronic resources and their coherent in-house control
- providing information support for the increasing patent activities within the Max Planck Society
- contributing to the clarification of legal and technical issues of access
- technical training, competent advice and user support for the scientists, librarians, administrators with regard to the most efficient use of the available information resources.
- maintenance, testing, and adaptation of existing software tools that facilitate the application of new information technologies by scientists of the Max Planck Society. The Service Component could provide the necessary technical support as well as identify and evaluate new ones as they become available.

The Center should be founded for a limited time-period of five years, and should be advised by a board of representatives of all three sections of the Max Planck Society. Its core group should have roughly the size of a „Nachwuchsgruppe“. Its foundation should go along with further support for and possibly a reorganization and enhancement of the present information management services of the Society, distributed over several locations. While the Development Component should closely interact with cutting-edge information management projects of the Institutes, the Service Component should integrate the existing services and enhance them by increased coordination and resources. Both components should be adminstratively integrated and be „customer-driven“. Failure to serve the information needs of MPG scientists should result in the rapid termination of the Center.