

The Impact of Electronic Publishing on the Academic Community

Session 6: Access to scientific data repositories

Introduction

Erich J. Neuhold

GMD-IPSI, Dolivostrasse 15, D-64293 Darmstadt, Germany

©Portland Press Ltd., 1997.

[Copyright Information](#)

With the advent of electronic production, storage, distribution and utilization of scientific data the role of all players is changing rapidly. In this session we investigate the changing role of data repositories of all kinds.

Traditionally there was a separation between: lending libraries (e.g. university libraries); collection libraries (e.g. the Library of Congress); and online databases (e.g. Cambridge Structural Database, chemical abstracts, etc.), with only limited co-operation between them.

However, these traditional roles are changing rapidly and there even exist opinions that their role will be so intermingled that only a single electronic service type will remain. Some of the questions that arise and that have been placed in front of the three contributors to this session are as follows.

What range of services will be offered by each of the three services? Will there be a one-stop supply for all the types of information kept either separately or redundantly in each of the three kinds of information service, or will each continue to offer specialized services?

What should be collected and kept? Should all information, e.g. that appearing on the Web, be kept for eternity or should only 'archival-value' information be kept? And what is 'archival value' for an archivist, for an engineer, for a scientist?

How can quality be ensured and communicated to the user? Should there be a strict reviewing process for information and should only that which passes the review be kept, or should everything be saved unreviewed and its value determined solely by the users?

How can durability and accessibility be ensured? What administrative and technological mechanisms have to be installed to keep the data readable and, even more important, searchable and presentable to future users? What does 'for all eternity' mean in this context?

What authentication, copyright, ownership, and pricing strategies and mechanisms will survive, have to be adapted or have to be invented? Will information become truly free (note, that is not the case now) or will pricing continue to exist because provision of the above three services will have costs attached to them? Who will carry those costs?

Who will be the customers? Will the currently overlapping but still quite distinctive clientele of the three services remain the same? Will the non-scientific community require, demand or even gain the right to access scientific data, e.g. for non-scientific publications or regulatory purposes?

What experiments, prototypes or actual solutions currently exist? Are we sufficiently confident that we understand the problems enough to develop solutions for the future lending libraries, collection libraries (archives) and online databases, or do we have to continue with technological, management and sociological research in order to determine the future form(s) of data repositories?

The following three papers will answer some of these questions and I will leave it to the reader to come to her/his own conclusions and to initiate actions where it seems fit. My own feeling is that many questions are currently being answered either theoretically, experimentally or in actual use, but that much more scientific and technological work is needed to arrive at stable, well-understood and widely accepted solutions.

So let us, on an interdisciplinary basis, team together to develop requirements, approaches and solutions which on one side are as generic as possible but on the other side are as discipline specific as needed. Time is passing rapidly and without action we may actually find ourselves starving for information despite the fact that there exists a practically unlimited (but unusable) information glut.

©Portland Press Ltd., 1997.

[Copyright Information](#)