



COMMISSION OF THE EUROPEAN COMMUNITIES

Brussels, 14.2.2007
COM(2007) 56 final

**COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN
PARLIAMENT, THE COUNCIL AND THE EUROPEAN ECONOMIC AND SOCIAL
COMMITTEE**

on scientific information in the digital age: access, dissemination and preservation

{SEC(2007)181}

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL AND THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE

on scientific information in the digital age: access, dissemination and preservation

1. INTRODUCTION

The present Communication comes from two policy strands, the **i2010 digital libraries initiative** and the **Community policy on research**. The digital libraries initiative aims to make information more accessible and usable in the digital environment. It follows up on a letter of 28 April 2005 by six Heads of State and Government asking the Commission to take necessary steps to improve access to Europe's cultural and scientific¹ heritage.

The Community policy on research looks to maximise the socio-economic benefits of research and development for the public good. The present Communication represents an initial step within a wider policy process addressing how the scientific publication system functions and what impact it has on research excellence. It comes at a strategic moment for European research with the launch of the Seventh Framework Programme (FP7) for 2007-2013 and the forthcoming Communication on developing the European Research Area (ERA).

This Communication's objective is to signal the importance of and launch a policy process on (a) access to and dissemination of scientific information², and (b) strategies for the preservation of scientific information across the Union. To this end, it announces a series of measures at European level and points to the need for a continuing policy debate.

These issues have a direct impact on Europe's capacity to compete through knowledge, a determining factor to reach the goals of the Lisbon agenda for competitiveness.

2. THE IMPORTANCE OF SCIENTIFIC INFORMATION

In order to become an increasingly competitive knowledge-based economy, Europe must improve the production of knowledge through research, its dissemination through education, and its application through innovation. All research builds on former work, and depends on scientists' possibilities to access and share scientific publications and research data. The rapid and widespread dissemination of research results can help accelerate innovation and avoid duplication of research efforts, although some delay for the first use by researchers or for commercial purposes can be justified. The system by which scientific information is published is pivotal for its certification and dissemination, and thus has a major impact on research funding policies and on the excellence of European research.

¹ In this Communication, the terms "scientific" or "science" refer to research activity in all scholarly subjects, including social sciences and the humanities.

² For the purposes of this Communication, 'scientific information' comprises publications and research data.

Public authorities fund around one third of European research³ and therefore have a clear interest in optimising the scientific information system. The stakes for the European Community are high: between 2007 and 2013, the Community will invest some €50 billion in FP7.

3. ACCESS TO AND DISSEMINATION OF SCIENTIFIC INFORMATION IN THE DIGITAL AGE

3.1. A system in transition: new markets, services and players

The rapidly increasing use of digital content in research and in the dissemination of knowledge is a main characteristic of modern science. The Internet makes instant access to and dissemination of scientific information possible and new information and communication tools offer innovative ways to add value. They have opened up new ways to use masses of data resulting from experiments and observations in the scientific process and to extract meaning from this data stored in repositories in combination with other scientific information resources. This leads to a "continuum" of the scientific information space from raw data to publications across different communities and countries.

Scientific journals traditionally hold a central role within the scientific information system. They are a vehicle for spreading research results and have a considerable impact on scientists' careers. The peer review process underpinning the selection of journal articles is its main quality control mechanism.

Technological change offers tremendous opportunities for Europe's scientific publishers. Over the past years, scientific publishers and other actors have made substantial investments in information technologies for online delivery, in the retro-digitisation of content, and in added-value services. About 90% of all science journals are now available online, in many cases via a subscription.

There are some 2 000 scientific journal publishers globally, producing about 1.4 million articles a year. Some 780 of these publishers are located in the EU, producing 49% of the total journal output. They employ some 36 000 persons directly in the EU and have a strong position in the world market.

An important recent trend has been the development of the open access movement, based on the viewpoint that access to publications and data should be improved in the Internet age. This movement aims to ensure immediate and free Internet access to research publications. A key milestone within this movement is the 2003 Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities.

According to the Berlin Declaration, Open Access publication requires that authors grant free access to their scientific contributions, as well as the possibility to use them, subject to proper attribution of authorship. Moreover, a complete version of the work and supplemental materials should be deposited in at least one online repository. 196 research institutions have signed the declaration, which is still open for signature.

The Open Access movement has resulted in publishers experimenting with the 'author pays' business model, which allows free online access to readers. This business model shifts

³ References backing up this statement and figures used in this document, as well as key definitions, can be found in the Commission staff working paper accompanying the present Communication.

publishing costs from the reader to the author, i.e. the author's institution or funding body. There are also increasing numbers of so-called hybrid journals offering both reader pay and author pay solutions. Another model currently being experimented with foresees a critical mass of journals in a specific area moving towards open access, under the sponsorship of a consortium of funding bodies. An example is the Sponsoring Consortium for Open Access Publishing in Particle Physics led by CERN, the European Organisation for Nuclear Research. At present, open access journals account for about 10% of total journal output.

A further development is the deposit of peer-reviewed and/or not yet peer-reviewed journal articles in freely accessible repositories organised by institution or discipline. In some cases, deposit occurs after an embargo period during which publishers can get a return on their investment (e.g. Cairn in France/Belgium). The length of the embargo period may vary between disciplines.

Some research funding agencies are active in developing policies recommending or mandating publication in open repositories of journal articles resulting from the research that they fund. Prominent examples are the Wellcome Trust and the National Institutes of Health. In the US Senate, a draft bill was recently introduced providing for federal agencies to develop public access policies.

The trends described have led to a debate on the scientific information system, with a focus on scientific journal articles. The key arguments used by stakeholders are summarised below.

Main arguments of researchers, research organisations, funding bodies, and libraries

- *Open access can increase the impact of scientific research and innovation through improved access to and rapid dissemination of research results.*
- *The Internet should bring the costs of scientific publications down, but journal prices have increased. This affects access to scientific information.*
- *The public purse pays for research, peer review (through reviewers' salaries), and journals (e.g. through library budgets). It is natural that public actors should request a better return on their investment.*

Main arguments of publishers

- *There is no access problem. Access to scientific information has never been better.*
- *Publishing has a cost. Publishers add considerable value to the research process by guaranteeing the quality of journal articles in the most efficient way possible.*
- *The publishing market is highly competitive and does not require public intervention. An ill-conceived intervention may lead to 'implosion' of the current system without offering a clear and viable alternative.*

3.2. Issues and challenges

Organisational issues

A shift in the type of publishing business model commonly used may entail unforeseen organisational consequences. For example, in an 'author pays' model, costs for accessing research results are shifted from one part of the public institution (the library) to another (e.g. university departments). This may lead to transitional costs or to a temporary gap in the accessibility of scientific information.

The emergence of increasing numbers of repositories containing not only peer-reviewed articles, but also working papers, PhD theses, research data, etc raises further issues.

An example of this integrated approach is the DARE programme in the Netherlands. Its objective is to provide networked free access to the academic output from all universities. The basic infrastructure currently includes more than 100.000 scientific reports and research articles, and in a later phase will include experimental or observational data, and other digital objects such as video- and audio-files.

Digital repositories promise new and integrated sources of information and are increasingly a strategic infrastructure in support of research. They require considerable organisational effort in relation to issues such as: who is responsible for depositing the material? How can the quality of repositories and of repository content (e.g. version management) be guaranteed? And how can repositories within Europe be linked to arrive at a critical mass of information?

Legal issues

When publishing scientific articles in journals authors normally assign their rights to publishers. Publishers then provide for the dissemination of these articles with an aim of ensuring a return on their investment. It has been suggested by researchers, funding bodies and libraries that current contractual practices can have a negative impact on access and dissemination and need to be reflected upon. The aim here is not the introduction of Community rules on copyright contract law – an area which has not been harmonised at Community level – but a reflection on the way in which authors exercise their rights in the digital environment.

In the case of research data, the intellectual property right (IPR) issue is different. While research data as such are not protected by IPRs, Directive 96/9/EC on the legal protection of databases⁴ protects efforts in organising research data. In this context, concerns have been raised about the impact of the *sui generis* right of protection for non-original databases on the accessibility of scientific research data.⁵

In case personal data is processed or disseminated in this context, access to and use of this data should respect the rules on the protection of personal data as laid out in EU Directives 95/46/EC and 2002/58/EC.⁶

Technical issues

Technological progress can greatly contribute to the accessibility and use of scientific information. For example, better search tools can help researchers find information and progress in new areas and collaborative tools can enhance the way in which researchers share information.

In order to link digital repositories and make them searchable, interoperability issues also need to be addressed systematically. In this respect, the use of open standards is crucial.

⁴ OJ L 77, 27.3.1996, p. 20.

⁵ See the 2005 evaluation report on the Directive by DG MARKT
http://ec.europa.eu/internal_market/copyright/docs/databases/evaluation_report_en.pdf.

⁶ OJ L 281, 23.11.1995, p. 31, and L 201, 31.7.2002, p.37.

Financial issues

Over the last twenty years, journal subscription prices have on average increased above inflation level - according to one study 4.5% per year above inflation - while there are considerable differences according to disciplines and journals. This has put publicly funded libraries, their main clients, under financial pressure and led to subscription cancellations in certain cases. This is particularly acute for less well-endowed institutions and in countries with lower income levels. Publishers argue that price increases are due to the growth in numbers of articles submitted and the increasing volume of journals, and that they are linked to a higher usage rate.

At the same time, the increase in research budgets has outpaced the funding of research result dissemination - today less than 1% of total European R&D-expenditure - including the available budgets of libraries. Publishers have responded by offering 'big deals' (bundling of journals for a discounted price) to libraries and to library consortia. These have benefited research organisations across Europe, but have also introduced the new problem of library budgets becoming inflexible through multi-year and relatively rigid contracts.

A further financial issue is that of value added tax (VAT) for digital products. Across Europe, digital journals are subject to standard rate VAT while paper journals benefit from a reduced rate. Thus the same content triggers a different VAT rate depending on the medium. Moreover, in view of the current rules concerning exemptions, public authorities and subsidies, public research institutes and libraries cannot deduct VAT costs. To address this situation, some Member States refund VAT for digital journal subscriptions to libraries.

4. PRESERVATION IN THE DIGITAL AGE

4.1. The problem

Long term preservation of digital material is a central problem for the information society, which features an exponentially growing and increasingly dynamic supply of information. Digital information is unstable due to rapid changes of hardware and software, and to the limited lifetime of storage devices. Information needs to be preserved in order to keep it readable and usable for the future. This challenge was signalled in the Communication 'i2010: digital libraries' in relation to Europe's digital cultural heritage.

The issue of preservation is relevant for publications and for research data. Preserving research data is essential to ensure traceability and repeatability of experiments. Moreover, research often depends on past observations, for example in the case of research on climate change. Sometimes research data collected in the past acquires contemporary relevance.

This was the case during the 'anthrax alerts' in the US and Europe. The British Library – one of the few places worldwide with comprehensive historic holdings of research information on anthrax – received numerous enquiries. Virtually no new research had been done on anthrax for forty years.⁷

There are currently no clear strategies in place across the Union for long term preservation and usability of digital scientific information. Existing national and European initiatives must

⁷ European Task Force Permanent Access, 'Permanent access to the records of science'.

be linked systematically. Preservation is also an area with considerable market potential (e.g. storage services) where Europe cannot afford to lag behind.

4.2. Issues and challenges

Organisational issues

The issue of preservation raises organisational questions. Who is responsible for preserving research data and the necessary software and hardware? What are the roles of research organisations and libraries? According to what criteria should the material to be preserved be selected? Moreover, a successful public preservation strategy requires good collaboration between public and private partners.

Examples of public-private partnerships for preservation purposes are the agreements between the Dutch National Library and publishers such as Reed Elsevier, Springer and Bio-Med Central.

Legal issues

Legal deposit, i.e. the obligation for content producers to make one or more copies of scientific materials available to a mandated deposit body, is a central issue for the preservation of digital scientific information. Member States have started to extend deposit arrangements to digital information, at different speeds and with different types of information covered. Nevertheless, the transition towards a digital environment may cause gaps in the intellectual record. A 2004 report of the Science and Technology Committee of the UK House of Commons signalled a gap of 60% in the deposit of electronically delivered publications due to delays in implementing legal deposit.⁸ In order to maximise the efficiency of the preservation process, digital information should be made available to mandated deposit bodies without technical protection against copying.

Technical issues

Technological advances can help to keep information accessible and usable. The goal is to reduce preservation costs and offer solutions for challenges such as the storage of large volumes of dynamic content. Upgrading of the supporting technical infrastructure would increase the capacity of research organisations to store information.

Financial issues

The cost of long term and sustainable preservation must be taken into account when setting up open repositories, but is often difficult to assess. Determining factors include the type and volume of information stored, the number of migrations needed, and envisaged use.

5. ACTIONS AT EUROPEAN LEVEL

5.1. Commission position

Initiatives leading to wider access to and dissemination of scientific information are necessary, especially with regard to journal articles and research data produced on the basis of public funding. With respect to journal articles, the Commission is observing and considering experiments with open access publishing.

⁸ HC Science and Technology Committee report 'Scientific Publications - Free for all?' - HC 399-1, July 2004, p. 93.

Fully publicly funded research data should in principle be accessible to all, in line with the 2004 OECD Ministerial Declaration on Access to Research Data from Public Funding.⁹

Moreover, the Commission draws particular attention to the need for clear strategies for the digital preservation of scientific information.

The Commission values the crucial role of all stakeholders in the scientific information system, and these stakeholders should be involved in any transformation process regarding access to, dissemination of and preservation of scientific information.

5.2. What has been done so far?

Member States and the Commission have started exploring issues of access, dissemination and preservation of scientific information, through **project funding** and by **launching a public debate with stakeholders**.

Examples of relevant projects co-funded under the Sixth Framework Programme (FP6) are CASPAR, DRIVER, and SEADATANET.

CASPAR looks at how to manage future access to and preserve scientific data. DRIVER focuses on ways to link repositories of scientific information. SEADATANET aims to develop a Pan-European Marine Data Management Infrastructure integrating the national marine data repositories.

The Commission has also begun to work with advisory groups and gather views from stakeholders, for example the High Level Group on Digital Libraries and the European Research Advisory Board (EURAB).

It has also financed a 'Study on the economic and technical evolution of the scientific publication markets in Europe',¹⁰ which was subject to a public consultation in 2006. Responses from these initiatives together with regular interaction with stakeholders have given the Commission valuable inputs.

At the **political level**, a Commission Recommendation on *the digitisation and online accessibility of cultural information and digital preservation* adopted on 24 August 2006 addresses the digital preservation issue.¹¹

5.3. Future actions managed by the European Commission

A. Access to Community funded research results

Within FP7, the Commission will take measures to promote better access to the publications resulting from the research it funds. In this context, project costs related to publishing, including open access publishing, will be eligible for a Community financial contribution. The Commission will encourage the research community to make use of this possibility.

The Commission also envisages, within specific programmes (e.g. the programmes managed by the European Research Council), to issue specific guidelines on the publication of articles in open repositories after an embargo period. This would be done on a sectorial basis, taking into account the specificity of the different scholarly and scientific disciplines.

⁹ Adopted in Paris on 30 January 2004. An OECD recommendation on this issue is currently in preparation.

¹⁰ http://ec.europa.eu/research/science-society/pdf/scientific-publication-study_en.pdf

¹¹ OJ L 236, 31.8.2006, p.28.

B. Co-funding of research infrastructures (in particular repositories) and projects

Within FP7 the Commission will intensify its activities regarding infrastructures relevant for access to scientific information, in particular by linking digital repositories at the European level. An amount of approximately €50 million will be made available to this end for the period 2007-2008 (some 20 million of which have been allocated for 2007).

In addition, an indicative amount of €25 million will be provided during this period (some 15 million of which during 2007) for research on digital preservation (in particular a network of Centres of competence for digital preservation) and on collaborative tools for using the content.

Within the *eContentplus* programme (2005-2008), €10 million has been earmarked to improve the accessibility and usability of scientific content, in particular addressing issues of interoperability and multilingual access.

C. Input for the future policy debate

To feed the debate and the policy process, the Commission will launch a study on the economic aspects of digital preservation to start in 2007. Moreover, through the Science in Society programme, the Commission will support research on the scientific publication system within the ERA and globally, for example on publication business models, dissemination strategies, and the connections between research excellence, scientific integrity and the scientific publication system.

When reviewing VAT-legislation, the Commission will critically examine issues relevant for scientific publications, such as the restrictions for public authorities or the exempt sectors on the recovery of VAT.

D. Policy co-ordination and policy debate with stakeholders

Discussions in the European Parliament and Council will contribute to a common understanding of access and dissemination issues at European level. In this respect, the Member States are invited to explore possible common strategies and to discuss the relevant issues and challenges – organisational, legal, technical and financial – highlighted in this Communication. Instruments such as ERA-NET and fora such as CREST and ESFRI could contribute to shaping the discussion.

The Commission will continue consultations with stakeholders in relevant expert and advisory groups such as the EIROforum, ESF, EURAB, and the High Level Group on Digital Libraries, taking into account the global dimension of the issue. It will organise a high-level conference on scientific publishing in the ERA in early 2007.

The Commission will encourage universities, research organisations, research funding bodies and scientific publishers to exchange information on good practices in relation to new access and dissemination models for scientific information.

Overview of actions

A. ACCESS TO COMMUNITY FUNDED RESEARCH RESULTS

- Costs for publishing, including open access publishing, defined as eligible costs in Community funded research projects.
- Specific guidelines to be issued, within specific programmes, on the publication of articles in open repositories.

B. CO-FUNDING THROUGH COMMUNITY PROGRAMMES

- Approximately €50 million for work on infrastructures, in particular digital repositories, in 2007-2008.
- Approximately €25 million for digital preservation and collaborative tools in 2007-2008.
- Approximately €10 million on access to and use of scientific information through *eContentplus* programme.

C. INPUT FOR THE FUTURE POLICY DEBATE

- Study on the economic aspects of digital preservation.
- Funding of research on publication business models and on the scientific publication system.

D. POLICY CO-ORDINATION AND DEBATE WITH STAKEHOLDERS

- Deliberations in the European Parliament and Council; further discussions with stakeholders.
- Exchange of good practices for new models of access to, dissemination of, and preservation of scientific information.

6. CONCLUSION

Access to, dissemination of, and preservation of scientific information are major challenges of the digital age. Success in each of these areas is of key importance for European information society and research policies. Different stakeholders in these fields have differing views on how to move towards improvements for access, dissemination and preservation.

Within this transition process from a print world to a digital world, the Commission will contribute to the debate among stakeholders and policy makers by encouraging experiments with new models that may improve access to and dissemination of scientific information, and by supporting the linkage of existing preservation initiatives at European level.

The Commission invites the European Parliament and Council to debate the relevant issues on the basis of the present Communication.