Moving from Restricted Dissemination of Publicly-Funded Knowledge to Open Knowledge Environments

CRICS8
Rio de Janeiro, BRazil
17 September 2008

by
Paul F. Uhlir
Director, Office of International Scientific and Technical Information Programs
The National Academies
Washington, DC
USA
puhlir@nas.edu

Revolution and Evolution in Scientific Communication

Comparison of some key characteristics of the print and digitally networked paradigms

PRINT	GLOBAL DIGITAL NETWORKS
(pre) Industrial Age	post-industrial Information Age
fixed, static	transformative, interactive
rigid	flexible, extensible
physical	"virtual"
local	global
linear	non-linear, asynchronous, with time/space collapsed
limited content and types	unlimited contents and multimedia
distribution difficult, slow	easy and immediate dissemination
copying cumbersome, not perfect	copying simple and identical
significant marginal distribution cost	zero marginal distribution cost
single user (or small group)	multiple, concurrent users/producers
centralized production	distributed production
slow knowledge diffusion	accelerated knowledge diffusion

Revolution and Evolution in Scientific Communication

Broad implications of excessive restrictions (economic, legal, technical) on access to and reuse of networked data and information from public sources:

1) **Disadvantage and marginalization of developing country or poor users.**
2) **Significant lost opportunity costs, and the related failure to capture maximum value from public investment in public data collection activities, including geospatial data.**
3) **Monopolization problems exacerbated in database markets, both public and private.**
4) **Higher transaction costs (not just cost of access).**
5) **Less effective international, inter-institutional, and interdisciplinary research cooperation using digital networks.**

***Openness* thus should be the default rule, subject only to legitimate and well-justified exceptions.**

Revolution and Evolution in Scientific Communication

What is an information commons?

- Digital data and information originating principally from government or publicly-funded sources;
- Made freely available for common use online;
- With the material in the public domain, or with only some rights reserved (using permissive licenses, such as Creative Commons, with “some rights reserved”), or with full intellectual property rights, but under open access conditions; and
- Typically organized thematically through an institutional mechanism.

Revolution and Evolution in Scientific Communication

Existing information commons models:

- **Open-source software movement (e.g., Linux and 10Ks of other programs worldwide, many of which originated in academia for research applications);**
- **Distributed Grid computing or *e*-science (e.g., Folding@Home, FightAIDS@Home, + >20 others in life sciences);**
- **Open data centers and archives (e.g., GenBank, Uniprot);**
- **Federated open data networks (e.g., World Data Center system, Global Biodiversity Information Facility;**
- **Open access journals (e.g., > 3500 scholarly journals, many in developing world—SciELO, BioMed Central);**
- **Open repositories for an institution's scholarly works (e.g., the Indian Institute for Science, + > 1000s/Ks? globally)**
- **Open repositories for publications in a specific subject area (e.g., the physics arXiv, CogPrints, PubMedCentral);**
- **Free university curricula online (e.g., the MIT OpenCourseWare);**
- **Discipline or applications commons or open knowledge environments (e.g., neurocommons, health commons).**

Revolution and Evolution in Scientific Communication

Advantages of information commons for science in academia:

- **Facilitates transfer of information North <-> South and South <-> South;**
- **Promotes capacity building in developing countries;**
- **Promotes interdisciplinary, inter-sector, inter-institutional, and international research cooperation;**
- Avoids duplication of research and promotes new research and new types of research;
- Reinforces open scientific inquiry and encourages diversity of analysis and opinion,
- Allows for the verification of previous results,
- Makes possible the testing of new or alternative hypotheses and methods of analysis;
- Facilitates the education of new researchers;
- Enables the exploration of topics not envisioned by the initial investigators;
- **Facilitates automated digital knowledge discovery and diffusion;**
- **Generally helps to increase the research potential of digital technologies and information, thereby providing greater returns from the public investment in research;**
- **Many other advantages and justifications**

Revolution and Evolution in Scientific Communication

Compelling reasons for placing government-generated data and information in the public domain or under open access conditions:

- ***Legal.*** **A government entity needs no legal incentives from exclusive property rights to create information. Both the activities that the government undertakes and the information produced by it in the course of those activities are a [global] public good.**
- ***Ethical.*** **The public has already paid for the production of the information. Burden of additional access fees falls disproportionately on the individuals least able to pay. Open access benefits the poor.**
- ***Governance.*** **Transparency of governance is undermined by restricting citizens from access to and use of public data and information. Restrictions on re-dissemination of public information, particularly of factual data, make governments less efficient and less accountable.**
- ***Socioeconomic.*** **Many economic and non-economic positive externalities. Network effects can be realized on an exponential basis through the open dissemination of data and information online—especially geospatial data. Conversely, the commercialization of public data and information on an exclusive basis produces de facto public monopolies that have inherent economic inefficiencies and are contrary to the public interest on other social, ethical, and good governance grounds.**

Revolution and Evolution in Scientific Communication

Elements of an open knowledge environment (OKE):

- Selection and articulation of the issue(s) to be addressed.
- Open access to the data and information that provide the inputs for the problem(s) and opportunities chosen for focus in the OKE.
- Open-source software and systems tools that support various types of networked collaboration.
- Strong project management.
- Effective social networking and outreach.

Revolution and Evolution in Scientific Communication

Examples of OKEs in health research:

- http://stke.sciencemag.org/
- http://sciencecommons.org/projects/data/background-briefing/
- http://sciencecommons.org/projects/healthcommons/

Revolution and Evolution in Scientific Communication

A vision for the future…

The restructuring of the print journal system through the formation of thematic OKEs in all universities

- Organized around a cluster of OA journals and databases.
- Managed by academic departments that integrate both domain discipline, computer science, and information science at one or more universities.
- Involving professors, students, and possibly professional societies.
- Focused on research, education, and development of new knowledge.

Revolution and Evolution in Scientific Communication

Barriers to creating information commons:

- Implementation and acceptance of new policy and institutional frameworks.
- Development of adequate incentives for participation at the individual, community, institutional, and governmental levels.
- Long-term financial sustainability of different information commons models.
- Effective technical and organizational approaches.
- In all cases, must balance with legitimate countervailing values and legal restrictions (protection of national security, privacy, confidentiality, and IPRs).

Revolution and Evolution in Scientific Communication

Additional background reading (all available freely online):

- ***Bits of Power: Issues in Global Access to Scientific Data (NAS, 1997)***
- ***The Role of S&T Data and Information in the Public Domain (NAS, 2003)***
- ***Reichman, J.H. and Paul F. Uhlir, "A Contractually Reconstructed Research Commons for Scientific Data in a Highly Protectionist Intellectual Property Environment, 66 Law & Contemporary Problems 315-462 (2003)***
- ***UNESCO Policy Guidelines for the Development and Promotion of Governmental Public Domain Information (2004)***
- ***Open Access and the Public Domain in Digital Data and Information for Science (NAS, 2004)***
- ***Strategies for Open Access to and Preservation of Scientific Data in China (NAS, 2006)***
- ***Strategies for Open and Permanent Access to Scientific Information in Latin America (CRIA, 2007)***