GeoData and the Law: A Research User Perspective

Paul F. Uhlir
Director, Board on Research Data and Information
National Academy of Sciences
Washington, DC
puhlir@nas.edu

GeoData and the Law:
A Research User Perspective

Comparison of some print and digitally networked paradigm characteristics:

PRINT PARADIGM	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
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 and integrated production
slow knowledge diffusion	accelerated knowledge diffusion
quasi private good	quasi public good

GeoData and the Law:
A Research User Perspective

What is a digital commons?

Digital data and information originating principally from government or publicly-funded sources;

- Made freely available for broad, common use online;
- With the material in the public domain, or with only some rights reserved (using common-use licenses, such as Creative Commons); and
- Typically organized thematically through an institutional mechanism.

GeoData and the Law:
A Research User Perspective

Advantages of open access to and unrestricted reuse of publicly generated or funded data and information on digital networks for science:

- **Promotes interdisciplinary, inter-institutional, and international research;**
- **Enables automated knowledge discovery;**
- **Avoids inefficiencies, including duplication of research;**
- 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;
- Supports studies on data collection methods and measurement;
- **Facilitates the education of new researchers;**
- **Promotes citizen scientists and serendipitous results,** enabling the exploration of topics not envisioned by the initial investigators and the primary research community;
- Permits the creation of new data sets when data from multiple sources are combined;
- **Promotes capacity building in developing countries and global research;**
- **Supports economic growth and social welfare; and**
- **Generally provides greater returns from public investments in research.**

GeoData and the Law:
A Research User Perspective

Compelling reasons for placing government-generated or government-funded data and information in the public domain or under common-use 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.**
- ***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.**
- ***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 and disadvantaged.**
- ***Political.*** **Transparency of governance is undermined by restricting citizens from access to and use of public data and information. Rights of freedom of expression and information are compromised by restrictions on re-dissemination of public information, particularly of factual data.**

GeoData and the Law:
A Research User Perspective

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

1) Higher research costs (monopolization of public goods, transaction costs)
2) Lost opportunity costs (automated knowledge discovery, failure to capture full benefits of public investments)
3) Barriers to innovation (new uses and serendipity limited)
4) Less effective scientific cooperation and education
5) Widening gap between OECD and developing countries

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

GeoData and the Law:
A Research User Perspective

Legitimate restrictions on public access to government data:

- National security and public safety
- Personal privacy
- Confidentiality
- Respecting proprietary rights of private-sector parties

GeoData and the Law:
A Research User Perspective

Key stakeholders in the development of scientific data access policies and laws—from the top down and the bottom up:

Top down policy development

- Governments
- Research funding agencies
- International and intergovernmental (scientific) organizations
- Learned societies (umbrella research community organization)

Bottom up policy development

- Universities and not-for-profit research institutes
- Industry research institutions
- Individual researchers
- General public

GeoData and the Law:
A Research User Perspective

Existing digital commons models and emerging open knowledge environments:

- Open-source software movement (e.g., Linux and 10Ks of other programs worldwide);
- Open data centers and archives (e.g., EROS Data Center, NERC UK);
- Federated open data networks (e.g., World Data Center System, GEOSS Data-CORE);
- Open access journals (e.g., > 5500 scholarly journals, in both more and less developed countries—i.e., SciELO, Bioline International);
- Open repositories for an institution's scholarly works (+ > 400 formally registered globally on Open DOAR, plus 1000s more not registered)
- Open repositories for publications in a specific subject area (e.g., the physics arXiv, CogPrints, PubMedCentral in US and UK);
- Free university curricula and lectures online (e.g., the MIT OpenCourseWare);
- E-government initiatives (Data.gov in US, many others worldwide); and
- Emerging discipline or applications commons, peer production of info, and integrated thematic open knowledge environments (e.g., virtual geo-observatories, wiki encyclopedias, open portals).

GeoData and the Law:
A Research User Perspective

Additional works by the presenter on this topic (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)*
- *Paul F. Uhlir, 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)*
- *Uhlir & Schröder, "Open Data for Global Science", Data Science Journal, CODATA, (2007).*
- *Uhlir, et al., Toward Implementation of GEOSS Data Sharing Principles, Journal of Space Law and Data Science Journal (2009)*