

Commons-based digital libraries

By

Anita Coleman

Institutional Address:

1515 E. First St.

School of Information Resources & Library Science

University of Arizona

Tucson, AZ 85719

Phone: +1 (520) 621-3565

Fax: +1 (520) 621-3279

Email: asc@u.arizona.edu

Commons –based digital libraries

Abstract

Commons-based digital libraries (CBDL) are an emerging phenomenon – they are digital libraries based on notions of common pool resource management and offering a sustainable and equitable vision for digital information management and use. The common-based digital library is first defined followed by the essential aspects of the framework. The metaphorical meanings and theories of libraries, repositories, and the commons are not included. Interested researchers are encouraged to contact the author.

Definition:

The commons-based digital library can be defined as global and trusted; its main purpose is the solution of large-scale, global or discipline-based problems through the provision of a common resource, digital information; collections and communities may be globally distributed but they are interdisciplinary and multi-disciplinary; they may be managed under common property or open access regimes; information resources include peer-reviewed research articles and diverse, digital objects of scholarship (datasets, learning objects); technical, cyber-infrastructure and emerging human information behaviors such as scholarly self-archiving and end-user classifying facilitate peer organization, multiple uses, reuses, new uses, and the peer production of knowledge.

CBDL are a socio-technical solution to one of the biggest problems that traditional and digital libraries face – how best to adapt and survive in an increasingly uncertain digital information future. Articulating a CBDL framework can also help us

find solutions for challenges to our current scholarly communication system especially those related to peer review, trust, and information quality. The quality of information is very difficult to measure and judge, as it is not only context-sensitive but includes a consideration of complex factors such as accessibility, accuracy, timeliness, and trust. For example, consider accessibility. The accessibility of scholarly information includes: physical access (either the print item is in a traditional library or the digital resource is technically accessible in one of several levels specified by web accessibility standards such as W3C (cite)), geographic access (many publications in the so-called South are not often available in the North), intellectual access (through activities such as indexing and inclusion in retrieval tools like the library catalog and bibliographic databases), open access (through open access archives and demolishing the twin barriers of economics and law), multi-lingual access (English-language translations of the whole or abstracts, etc.), and perpetual access (the materials is available forever). Finally, CBDL are a different way of typing digital libraries from the usual digital library research versus practitioner or hybrid library perspectives and categories (Lesk, 1998, 2004; Borgman, 2000; Chowdhury and Chowdhury, 2003; Marcum and George, 2006). CBDL cannot and do not emerge out of thin air; they are progressions from prior frameworks such as the community-based digital libraries and scholarship repositories of which DLESE, the Digital Library for Earth System Education, ArXiv, the High Energy Physics repository, and the California Digital Library, are excellent forerunners and examples.

Framework:

Three of the components of the CBDL framework to be investigated further are:

1) Nature of digital information resources

CBDL goods are the many types of scholarly information products; these can range from peer reviewed published postprints to working papers but the key aspects of the nature of the information resource in this context revolves around questions of complexity, simplicity, subtractability, and exclusion. Digital information is usually not a subtractable good; that is, its consumption often actually increases use and production (more information is created); neither does digital use appear to detract from its availability or accessibility to another because it is easy to simultaneously share with larger numbers unlike a physical book. It is also more difficult to exclude beneficiaries to digital information. Are these universal and applicable to all information resources? What are limitations and when do the limitations start to apply?

2) Property regime of the digital information resource

The classic definition of an open access property regime is that no one has the right to limit use; in common property regimes, on the other hand, members have the right to set limits on use such as excluding nonmembers from using that resource. Digital scholarly information is typically owned by a number of different stakeholders; technologies such as digital repositories allow authors to retain ownership of copyright but transfer distribution rights and request attribution. CBDL have technologies and policies that depend on both common property and open access regimes. The nature of information use and the size of the field of active use (for example, number of disciplinary researchers) are key determinants of the appropriate regime for the various information resources and users. Does the user want to only read the resource? Do they cite the

resource? Where does citation fall in the continuum of use? Or might they indulge in further extractive uses such as quoting a snippet from a research article or using the research instrument in its entirety in their own research? What are synchronous and asynchronous interactions between users and data? A better articulation of digital information uses that are grounded in actual use in various contexts, technology affordances and barriers is needed; this is something traditional libraries have found difficult to do and digital libraries have not solved. Understanding disciplinary differences and how hybrid research communities form, separate, and renew or are reborn in a new configuration is critical. The development of sustainable aggregation, federation, and trusted services, that facilitate transformational uses depend on clearer articulation of these issues.

3) Defining access and distinguishing it from limits on amount of access and the nature of the use

Assuming an open access property regime, that is all users have open access to the material, may be a fundamental requirement. But the nature of the access itself – universal, granular or at other levels should be distinguished. What is the nature of the problems to be solved by the aggregation of information and services developed? This means that certain types of material (research datasets or protocols or learning objects such as syllabi, for instance) may be more amenable to multiple use/re-use, modification and type of distribution and computational services (deposit and metadata creation versus user annotations for instance), required. Understanding the nature of open access property regimes and their impact on problems that need to be solved, the forms and

formats of information necessary to effectively limit and regulate use are important issues for investigation.

Conclusion

The question of whether some forms of information resources such as peer reviewed research articles or the outputs of publicly funded research are analogous to property such as land is not a new one. Even in general property theory, the view that greater propertization leads to greater efficiency has been challenged. Information in general, and the majority of specific types of information resources (instantiations of information) have also long been shown to be different from other common property resources; information shows no efficiency gains when the standard economic model of a private property system is enforced. Restriction (limiting access) and enforcement (ensuring that only authorized users have access) also have transaction costs; additionally, they reduce the value of the information through lack of use. Law, technology, and economic models do not handle the problem of insecure, overlapping, and unenforced property rights very well; these tend to be major defining characteristics of most scholarly digital information resources. Publicly funded and academic research may better fit the open access property regime in which the resource is so abundant, mobile, and dispersed that authors and users can rationally decide not to engage in acts of exclusion at all. Unfortunately, we don't know a lot about such property regimes. Using the CBDL framework will identify taxonomic distinctions that are grounded not merely in the technology but in the nature of information resource uses and strategic forms of user behavior.

Acknowledgments

Thanks to Blaise Cronin for very helpful comments on a very early draft! Thanks to the faculty at Indiana University who helped me develop some of these ideas by asking lots of questions. Thanks also to Heather Morrison for helping me refine the definition.

References and Notes

Arms, W. Y. 2000. Digital libraries. Cambridge, Mass.: MIT Press.

Benkler, Y. (2002) Coase's Penguin, or, Linux and *The Nature of the Firm*. Yale Law Journal 112. Retrieved from Lexis Nexis Academic Universe (subscription database)

Borgman, C.L. (2000). From Gutenberg to the global information infrastructure: access to information in the networked world. Cambridge, Mass.: MIT Press.

Chowdhury, G.G. and Chowdhury, S. (2003). Introduction to digital libraries. London: Facet publishing.

Conservation Commons. (2006). <http://www.conservationcommons.org/>

Creative Commons. (2006). <http://www.creativecommons.org/>

Cronin, B., Shaw, D. and La Barre, K. (2003). A cast of thousands: Co-authorship and sub-authorship collaboration as manifested in the scholarly literature of psychology and philosophy. *Journal of the American Society for Information Science* 54 (9): 855-879.

Crowe, B. (1969). The tragedy of the commons revisited. *Science* 166: 1103-1107.

Davenport, E., Cronin, B. The citation network as a prototype for representing trust in virtual environments. In Cronin B., Atkins H. B. (eds.). *The Web of Knowledge: a Festschrift in Honor of Eugene Garfield*. Information Today Inc. & The American Society for Information Science: Medford, NJ, 2000

Ginsparg, P. (1996). Winners and losers in the global research village. In Conference held at UNESCO HQ, Paris, 19-23 Feb 1996, during session *Scientist's View of Electronic Publishing and Issues Raised*, Wed 21 Feb 1996.

<http://people.ccmr.cornell.edu/~ginsparg/blurbs/pg96unesco.html> (Accessed February 10,

Goner, E.C.K. 1966. *Common land and inclosure*. 2nd ed. London: Cass.

Hardin, G. (1968). The tragedy of the commons. *Science* 162, 1243-1248.

Hardin, G. (1998). Extensions of "The tragedy of the commons." *Science* 280: 682-683.

Harter, S. (1997). Scholarly communication and the digital library. *Journal of Digital Information* 1 (1). <http://jodi.tamu.edu/Articles/v01/i01/Harter/> (Accessed February 10, 2006).

Hess, C. and Ostrom, E. (2003). Ideas, artifacts, and facilities: Information as a common-pool resource. *Law and Contemporary Problems* 66 (Winter/Spring, 1&2): 111- 145.

Hess, Charlotte. (2000). "Is There Anything New Under the Sun? A Discussion and Survey of Studies on New Commons and the Internet." Presented at "Constituting the Commons: Crafting Sustainable Commons in the New Millenium", the Eighth Conference of the International Association for the Study of Common Property, Bloomington, Indiana, USA, May 31-June 4.

Kling, R., McKim, G. & King, A. (2003) A Bit More to IT: Scientific Multiple Media Communication Forums as Socio-technical Interaction Networks. *Journal of the American Society for Information Science*, 54(1), 47-67.

Kling, R. & McKim, G. (2000). Not just a matter of time: Field differences in the shaping of electronic media in supporting scientific communication. *Journal of the American Society for Information Science* 51(14).

<http://xxx.lanl.gov/ftp/cs/papers/9909/9909008.pdf>

Lesk, M. (1997). *Practical digital libraries: Books, bytes, and bucks*. San Francisco, CA.: Morgan Kaufman.

Lesk, M. (2004). *Understanding digital libraries*. 2nd ed. San Francisco, CA: Morgan Kaufman.

Lloyd, William Forster. 1833. Two lectures on the checks to population, delivered before the University of Oxford, in Michaelmas Term 1832. In Hardin, G. and Baden. *Managing the Commons*. pp. 8-15.

Lynch, Clifford. (2003). *Institutional Repositories: Essential infrastructure for scholarship in the digital age*. *ARL Bimonthly Report 226*: February.
<http://www.arl.org/newsltr/226/ir.html>.(Accessed February 10, 2006).

Mannheim, A. (2000). When did peer review become anonymous? *Aslib Proceedings* 52 (8): 273-276.

Marcum, Deanna and George, G. (2006). *Digital library development: The view from Kanazawa*. Westport, Conn: Libraries Unlimited.

Neelameghan, A. 1997. International cooperation and assistance. *World Information Report*. Paris: UNESCO. <Http://www.unesco.org/cii/wirerpt/vers-web.htm>. Chapter 27, pp. 361–380.

Neelameghan, A., and Tocatlian, J. 1985. International cooperation in information systems and services. *Journal of the American Society for Information Science* 36 (3): 153–163.

Ginsparg, P. (1996). Winners and losers in the global research village.
<http://people.ccmr.cornell.edu/~ginsparg/blurb/pg96unesco.html#current>

RLG. (2002). Trusted digital repositories: Attributes and responsibilities. (An RLG-OCLC Report). Mountain View: RLG.

Robbin, A. (2005). Rob Kling in search of one good theory: The origins of computerization movements. Presentation at CRITO, UCI.

Schatz, B., Bishop, A., Mischo, W., and Hardin, J. (1994). Digital library infrastructure for a university engineering community. In *Digital Libraries 1994. Proceedings of the First Annual Conference on the Theory and Practice of Digital Libraries*, College Station, Texas, June 19-21, 1994.
<http://www.csd.tamu.edu/csd/DL94/paper/schatz.html>. (Accessed February 10, 2006).

Tennant, R. (2002). Institutional Repositories. *Library Journal*, September 15.
<http://www.libraryjournal.com/index.asp?layout=articlePrint&articleID=CA242297>
(Accessed February 10, 2006).

Tennant, R. (2005). The Open Content Alliance. *Library Journal*,
<http://www.libraryjournal.com/article/CA6289918.html> (Accessed February 13, 2006).

Van House, N. (2002). Trust and epistemic communities in biodiversity data sharing. In
International Conference on Digital Libraries: Proceedings of the 2nd Joint ACM/IEEE
Digital Libraries Conference, Portland, Oregon. (pp. 231-239). New York: ACM Press.