

Response to Request for Information "Public Access to Peer-Reviewed Scholarly Publications from Federally Funded Research", November 2011

January 12, 2012

Clifford Lynch
Executive Director
Coalition for Networked Information
Cliff@cni.org

I am pleased to have the opportunity to submit comments to this request for information on "Public Access to Peer-Reviewed Scholarly Publications from Federally Funded Research" on behalf of the Coalition for Networked Information (CNI). CNI is a membership organization consisting of some 200 organizations, primarily but far from exclusively universities, who share a common commitment to advancing the intelligent use of information technology and digital content in support of scholarship. You can find more information on CNI at www.cni.org.

I want to be clear that while these comments are certainly informed by discussions with CNI's member organizations, they should not be viewed as representing the position of any specific member of CNI.

In general, CNI supports the analysis in the response to this call for comments already submitted by Prudence Adler on behalf of the Association of Research Libraries (ARL) on January 8, 2012, and available at <http://www.arl.org/pp/access/accessfunded/rfi-access-to-pubs-8jan2012.shtml>. It is clearly time to extend public-access policies to all federally funded research.

I want to supplement the ARL analysis with two additional points that speak to areas where CNI has focused some specific attention and expertise in recent years.

I believe that scholarly and scientific norms, as well as sound public policy, support the practice that the underlying data supporting published results need to be publicly available to facilitate replication and reproduction of those results. Also, their availability is important for additional scholarly analysis and re-use. These arguments have been widely presented in scholarly journals, governmental and scholarly policy reports, and studies by the National Academies. As journal articles and other forms of scholarly publication begin to move beyond the constraints of the historic printed page and exploit the affordances of the digital environment in which scholarly publications now exist, the boundaries between publications and underlying data will rapidly become much more fluid. A clear understanding about public access to publications will facilitate access to underlying data (as well as the understanding and reuse of this

data); similarly, barriers to public access to publications will create obstacles to public access to the underlying data.

My second additional point deals with the changing nature and use of the scholarly literature. For a number of reasons not just the size but the rate of growth of the scholarly literature is increasing steadily. It is increasingly unrealistic for unassisted human scholars to cope with this rate of growth; there's a new paper published every minute or two, every day of the week, every week of the year. We need to be able to apply information technology in more sophisticated ways to help scholars to deal with this flood of information; literally, to compute on the literature. As long as the vast majority of scholarly literature is scattered across an archipelago of proprietary, access-restricted silos, development and deployment of these computational tools to manage, navigate and mine the scholarly literature will be largely impossible. Public access to the publications from federally funded research - if access is appropriately defined to include these types of computational access - will substantially help in creating an environment that will facilitate the development and adoption of these computationally assisted discovery technologies, to the advantage of both scholarship and commercial exploitation of the body of scholarly knowledge.

I conclude with a few citations to work that explores these two points in more detail. Clifford A. Lynch, "Jim Gray's Fourth Paradigm and the Construction of the Scientific Record", *The Fourth Paradigm: Data-Intensive Scientific Discovery*, Tony Hey, Stewart Tinsley, and Kirstin Tolle (Eds.), (Redmond, WA: Microsoft Research, 2009), pp. 177-183. Online at <http://research.microsoft.com>

Clifford A. Lynch, "The Shape of the Scientific Article in the Developing Cyberinfrastructure", *CT Watch* 3:3 (August 2007), pp. 5-11. Online at www.ctwatch.org.

Clifford A. Lynch, "Open Computation: Beyond Human-Reader-Centric Views of Scholarly Literatures", *Open Access: Key Strategic, Technical and Economic Aspects*, Neil Jacobs (Ed.), (Oxford UK: Chandos Publishing, 2006), pp. 185-193. Online at www.cni.org.