



## ASSOCIATION OF RESEARCH LIBRARIES

### **Establish a Universal, Open Library or Digital Data Commons**

*"The Obama plan represents not new public works but, rather, investments that will work for the American public. Investments to build the classrooms, laboratories and libraries our children need to meet 21st-century educational challenges." Lawrence Summers, Washington Post, December 28, 2008.*

#### **Summary:**

Deepening our understanding of our Nation and its culture and history, advancing scientific discovery, tackling environmental, economic issues and more, all depend on scientists, researchers, students, scholars, and members of the public accessing our Nation's cultural, historical and scientific assets. A large-scale initiative to digitize and preserve the public domain collections of library, governmental, and cultural memory organizations will support research, teaching and learning at all levels, will help stem the current economic crisis by equipping and employing workers in every state with 21st Century skills, and it will lay a foundation for innovation and national competitiveness in the decades ahead. The goal is to establish a universal, open library or a digital data commons.

#### **The Need:**

New investments in cyber and information infrastructure are critical components to advancing science and education, spurring innovation, sparking civic engagement, and tackling key issues of importance to the Nation. Recent reports such as the National Academies report, *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*, reflect the pressing need to ensure an environment that is conducive to enabling the United States to meet the global challenges of the 21<sup>st</sup> century. This means that researchers, students, and the public must be empowered by having the full array of information resources needed to make contributions in all arenas. Broad distribution of information and research enables scientists, including citizen scientists and university researchers, to build upon it and approach problems with new perspectives. It permits educators and students to have access to needed resources previously unavailable without regard for geographic location or financial limitations. And it gives members of the public the ability to understand their culture and history in new, meaningful ways.

We have learned from the scientific community that many discoveries result from building on prior studies. For example, the discovery of the structure of DNA, the development of penicillin, and the development of radiation treatment for cancer patients all stemmed from researchers building on the work of others. It is time to extend discovery and access well beyond current bounds to foster new educational applications and encourage greater civic engagement.

For example, educators and students at all levels would benefit from web accessible educational materials. Today, the vast majority of these public domain resources remain inaccessible because they are in print or microform – formats no longer preferred nor beneficial given the nature of the technology -- by users of information. These resources also are distributed geographically throughout the United States. They include governmental, historical, cultural and scientific information resources. A digital, universal library or digital data commons could include resources that span collections of research libraries, museums, and other cultural memory organizations. It will open the doors of these institutions to the public without regard to where they live. Such a commons could also include new educational and creative works specifically developed to complement existing public domain works in the commons. Examples of resources include:

- Digitized full-text scientific and technical R&D reports dating from 1940s to present.
- A wealth of resources such as non-text scientific research data including images, audio, video, and numeric data.
- Legacy collections of government agencies spanning over 200 years and covering virtually every facet of U.S. history, government, policy and administration.
- Course materials, including non-print media, developed throughout the Nation. The Massachusetts Institute of Technology has already demonstrated the strong demand for its faculty's lectures and related resources with their Open Courseware project.

As selected library collections have become digitized and publicly available, there have been new and innovative applications in the classroom. And importantly, particularly in this economic climate, the broad availability of these assets acts as a "leveler" by expanding the number of potential users without regard for institutional affiliation throughout the country, indeed the world. This, in turn, will allow for greater sharing of information and the spurring of education.

Investment in such an initiative requires collaboration and cooperation amongst participating libraries and institutions in every State to ensure the development of an effective, distributed network. Moreover, the extended deployment of broadband will be essential as it will not only provide needed connectivity for users but it will be key to harnessing the power of the distributed library and related expertise and resources, and be the conduit for new content.

### **Costs and Opportunity:**

Despite significant technological and fiscal investments by Government, industry and academe, barriers still remain to achieving the full potential of our information infrastructure. The core capacity to bring library, cultural memory organization and governmental collections online exists now, ready to be scaled up. Libraries have long understood the immense value in making their collections broadly available. Many libraries have been selectively scanning their print collections for more than a decade but funds have been scarce to make

dramatic headway or undertake comprehensive library digitization. Recently Google and Microsoft have made investments in scanning books but commercial and legal requirements underlying their investments, limit full and free access.

There have been other digitization projects that provide data on how to scale and undertake such an endeavor. For example, in collaboration with libraries and aided by contributions from foundations, the non-profit Internet Archive has put a million printed books online. It currently operates 18 U.S. scanning centers with capacity to employ 600 people. With an injection of funds to numerous institutions and organizations, the rate of digitization could be quickly ramped up. It is estimated that with a graduated order, 10,000 individuals in all 50 states could be trained and put to work scanning books, manuscripts, journals, and other public domain materials. With the inclusion of other resources, these numbers would increase.

To initiate a shovel ready, universal, open library or digital data commons of 10 million books would require an estimated \$300 million. To extend this initiative well beyond books would require additional funds.

The Institute of Museum and Library Services (IMLS) and the National Endowment for the Humanities (NEH) are two agencies that would provide many valuable strengths and contributions to such an initiative. Via the National Leadership Grants Program, IMLS has funded digitization projects and worked effectively at building distributed, collaborated projects within the library and museum communities. NEH has been funding preservation and access projects in the humanities and brings extremely valuable experience to bear on preservation issues.

**Conclusion:**

A large-scale initiative to digitize public domain collections meets just about any test of an effective response to the mounting problems that challenge the United States. Beyond retraining workers with new, valuable skill sets and putting them to work, this initiative will bring high-quality public domain resources into every home, school, community college, university and workplace. It will give businesses, state and local governments, and jobseekers needed resources and will enrich education at all levels by bringing the world's collective knowledge to parents, teachers, and students. Finally, these scientific, cultural and historical assets will provide much needed content to the extended deployment of broadband throughout the country. Above all, these online, high value intellectual resources will remain available permanently to the Nation as research libraries will provide long-term preservation and access to the digitized content.

January 2009

---

Prue Adler  
[prue@arl.org](mailto:prue@arl.org)  
202-296-2296

---

Charles Lowry  
[charles@arl.org](mailto:charles@arl.org)  
202-296-2296

<http://www.arl.org/pp/index.shtml>