Tue 11/8/2011 12:51 PM

Request for Information: Public Access to Digital Data Resulting From Federally Funded Scientific Research

Preservation, Discoverability, and Access

(1) What specific Federal policies would encourage public access to and the preservation of broadly valuable digital data resulting from federally funded scientific research, to grow the U.S. economy and improve the productivity of the American scientific enterprise?

Data and associated software should have an entry in bibliographic listings together with the citation of the corresponding article. This should be enforced through the journals that publish the article.

(2) What specific steps can be taken to protect the intellectual property interests of publishers, scientists, Federal agencies, and other stakeholders, with respect to any existing or proposed policies for encouraging public access to and preservation of digital data resulting from federally funded scientific research?

(3) How could Federal agencies take into account inherent differences between scientific disciplines and different types of digital data when developing policies on the management of data?

(4) How could agency policies consider differences in the relative costs and benefits of long-term stewardship and dissemination of different types of data resulting from federally funded research?

(5) How can stakeholders (e.g., research communities, universities, research institutions, libraries, scientific publishers) best contribute to the implementation of data management plans?

(6) How could funding mechanisms be improved to better address the real costs of preserving and making digital data accessible?

(7) What approaches could agencies take to measure, verify, and improve compliance with Federal data stewardship and access policies for scientific research? How can the burden of compliance and verification be minimized?

(8) What additional steps could agencies take to stimulate innovative use of publicly accessible research data in new and existing markets and industries to create jobs and grow the economy?

(9) What mechanisms could be developed to assure that those who produced the data are given appropriate attribution and credit when secondary results are reported?

Data and associated software should have an entry in bibliographic listings together with the citation of the corresponding article. This should be enforced through the journals that publish the article.

Standards for Interoperability, Re-Use and Re-Purposing

(10) What digital data standards would enable interoperability, reuse, and repurposing of digital scientific data? For example, MIAME (minimum information about a microarray experiment; see Brazma et al., 2001, Nature Genetics 29, 371) is an example of a community-driven data standards effort.

For computational modeling software, there are already developed standards (SBML, NeuroML, and related).

(11) What are other examples of standards development processes that were successful in producing effective standards and what characteristics of the process made these efforts successful?

(12) How could Federal agencies promote effective coordination on digital data standards with other nations and international communities?

(13) What policies, practices, and standards are needed to support linking between publications and associated data?

Data and associated software should have an entry in bibliographic listings together with the citation of the corresponding article. This should be enforced through the journals that publish the article. This could be enforced through the PubMed database for example, as a requisite for correct citation in the database.

Response to this RFI is voluntary. Responders are free to address any or all the above items, as well as provide additional information that they think is relevant to developing policies consistent with increased preservation and dissemination of broadly useful digital data resulting from federally funded research. Please note that the Government will not pay for response preparation or for the use of any information contained in the response.

German Cavelier PhD, Office of Technology Development and Coordination

National Institute of Mental Health