January 4, 2012

Mr. Ted Wackler
Deputy Chief of Staff
Office of Science and Technology Policy
725 17th Street
Washington, DC 20502

Re: Public Access to Digital Data Resulting From Federally Funded Scientific Research (76 FR 68517)

Dear Mr. Wackler:

I am writing on behalf of Oxford University Press (OUP) in response to OSTP's Request for Information regarding "Public Access to Digital Data Resulting from Federally Funded Scientific Research".

The world's largest university press, Oxford University Press is an international publisher of scholarly and educational material with offices across the globe including major centers in New York City and in Cary, NC. OUP furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide and publishes over 270 peer reviewed scholarly journals (most of which are published in partnership with learned societies) and more than 2000 research monographs per year .

Preservation, Discoverability, and Access
As a scholarly publisher working closely with the academic community, OUP supports efforts to make data more available and more discoverable, and to ensure that appropriate data is preserved and curated for future use.

The first step is the creation of an environment and infrastructure where investigators are able and encouraged to deposit their data. While data management is built-in to the experimental design of so called 'big science', data management for the remaining 90% of scholarly endeavor could be improved. To this end, federal agencies should work with stakeholders including researchers, learned societies, library groups, and publishers and:

- set aside funds to support the establishment of subject-specific archives for data where they do not already exist
- develop policies and guidance on what data should be deposited
- develop policies which encourage researchers to deposit data and which balance the needs of investigators to gain credit by utilizing the data that they produce with the opportunities to advance knowledge by sharing data with other researchers and with the public
- encourage the development of metadata standards that can be used to describe data in order to improve discoverability
- develop preservation criteria regarding which data should continue to be preserved and curated

Standards for Interoperability, Re-Use and Repurposing
We see the key issues with respect to interoperability and use of primary research data as the development of:

- standards for the bi-directional linking between primary data and the peer reviewed research literature, and data citation
- standards for the acknowledgement of the use and re-use of data
- clear rules on modification of source data and how this modification is described
- security protection protocols to guard against unauthorized modification, damage, or deletion

OUP notes that standards for data stewardship are currently at an embryonic stage but that the following are examples of good projects / initiatives that we respectfully suggest Federal agencies should engage with:

APARSEN (www.alliancepermanentaccess.org/index.php/currentprojects/aparsen)
CASPAR (www.casparpreserves.eu)
CoData (www.codata.org)
DataCite (datacite.org)
DCC (www.dcc.ac.uk/)
DRYAD (www.datadryad.org)
nestor (www.langzeitarchivierung.de)
NISO/NFAIS Supplemental Journal Materials Working Group (www.niso.org/workrooms/supplemental)
OAIS (public.ccsds.org/publications/archive/650x0b1.pdf)
Opportunities for Data Exchange (www.ode-project.eu)
PARSE.insight (www.parse-insight.eu)
Planets (www.planetsproject.eu)
SHAMAN (www.shaman-ip.eu)

Yours sincerely,

Niko Pfund
President and Academic Publisher
Oxford University Press