A memo dated 22 Feb. 2013 of this year from John Holdren, President Obama’s director of the Office of Science and Technology Policy (OSTP), presented the framework to guide federal agencies charged with increasing access to the results of federally funded scientific research, including publications and data. While “open access” publication is well known to agriculture as part of the business models of many peer-reviewed journals, open access data is not only a novel concept to many agricultural researchers, but one for which we are not presently well equipped. Achieving functional open access to all agricultural research data will require the development and implementation of significant, novel infrastructure including standards, policies for sharing, embargoing, preservation and curation, workflows and educational modules to facilitate data preparation, and repositories tailored to the unique object attributes of data.

According to the OSTP memo, the open access data premise is simple. Federally funded research catalyzes innovation and breakthroughs that stimulate economic growth; research data should be viewed as an asset underpinning progress in health, energy, environment, and agriculture. Thus, policies and infrastructure that create functional data repositories would enable data preservation and subsequent reuse, maximizing the impact and accountability of federal research investment. Most agencies funding research of ASA, CSSA, and SSSA members are expected to create a public-access plan for data in the coming year, including USDA.

What is considered “data” and which data falls under this mandate? The OSTP memo defines data “as the digital recorded factual material commonly accepted in the scientific community as necessary to validate research findings, including datasets used to support scholarly publications...” Digitally formatted scientific data resulting from research supported wholly or in part by federal funding should be stored and publicly accessible to search, retrieve, and analyze. Specifically excluded as data are “laboratory notebooks, preliminary analyses, drafts of scientific papers, plans for future research, peer-review reports, communications with colleagues, or physical objects, such as laboratory specimens.”

Each agency is charged with providing an assessment of long-term needs for data preservation that is inclusive of all of the disciplines it supports. Further, the memo specifically encourages plans that complement current efforts already ongoing among public- and private-sector stakeholders. As detailed in the memo, data access infrastructure and policies being developed in response to this directive are expected to:

- Maximize data access free of charge to the public, protecting confidentiality and personal privacy, recognizing proprietary interests, and preserving the balance between the relative value of long-term data preservation and access and the associated cost and administrative burden. Use of publicly available databases, where possible, is encouraged.
- Improve data preservation and access through cooperation/collaboration, including the formation of public–private partnerships.