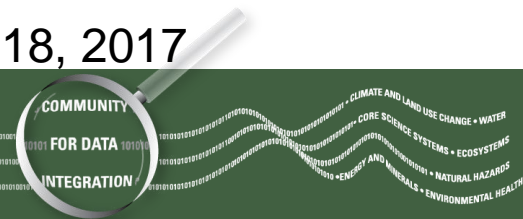


Challenges to and progress towards open science from both federal and community-driven perspectives

Lindsay Powers
US Geological Survey
National Geological and Geophysical Data Preservation Program

National Academies of Sciences, September 18, 2017



- One day, back in February 2013, Office of Science, Technology and Policy released a memorandum.
- The subject line said, “Increasing Access to the Results of Federally Funded Scientific Research” ...
- ...and it asked that USGS produce a plan.



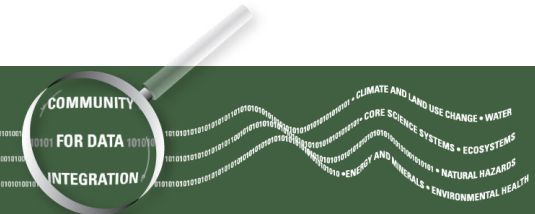
EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY
WASHINGTON, D.C. 20502

February 22, 2013

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: John P. Holdren
Director

SUBJECT: Increasing Access to the Results of Federally Funded Scientific Research



The Memo was part of a suite of the Open Data Initiatives – 2013

Executive Order

- Increasing Public Access to the Results of Federally Funded Scientific Research Feb. 22, 2013

Executive Order

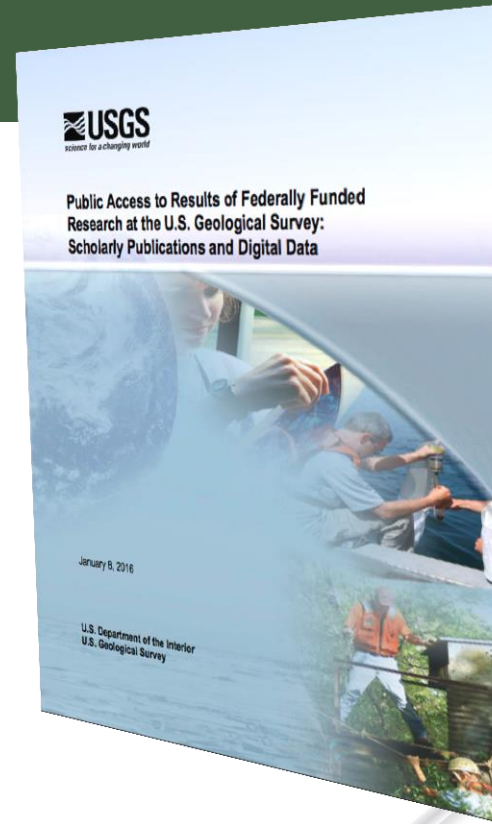
- Making Open and Machine Readable: the New Default for Government Information May 9, 2013

M-13-13 Open Data Policy

- Managing Information as an Asset

“Public Access to Results of Federally Funded Research at the U.S. Geological Survey: Scholarly Publications and Digital Data”

- outlines a framework for activities to increase public access to scholarly publications and digital scientific data resulting from research funded by the USGS



Components of an effective USGS Public Access Plan

OVERARCHING GOAL: electronic copy of final accepted manuscript or final publication of record AND supporting data is available free of charge for public access not more than 12 months after publication date.

Data Management Plans

Formal **Data Management Plans** must accompany all new research proposals

IPDS

(Information Product Data System) Full-text of all USGS final manuscripts, and USGS-funded publications by scientists extramural to USGS, will reside in **IPDS in a dark archive**

Pubs Warehouse

Catalog record of all USGS series pubs and externally published pubs in the USGS **Publications Warehouse.**

CHORUS

CHORUS (the Clearinghouse for the Open Research of the United States) will provide USGS with notification that extramural publications have been published on non-USGS websites.

More components of an effective USGS Public Access Plan

Review and Approval

Data undergoes QA/QC, formal description, **review and approval before release**

USGS repositories

Data are submitted to a **USGS repository**

All intramural data and publications are stored in **trusted digital repositories**

Science Data Catalog

USGS Data Metadata are indexed in **USGS Science Data Catalog**

Digital Object Identifiers

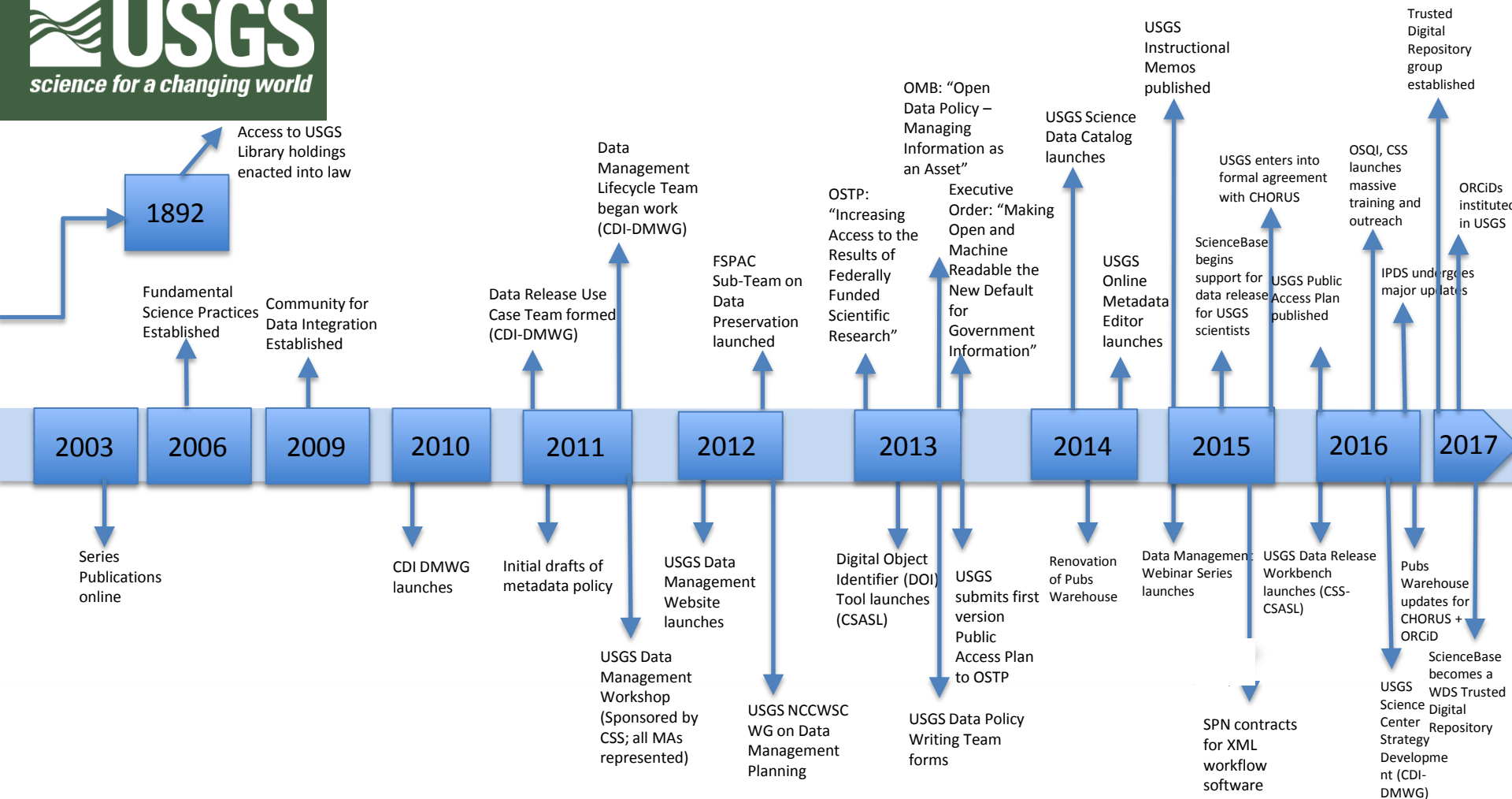
Digital Object Identifiers assigned for data (USGS DOI Tool to access DataCite) and publications (IPDS assigns CrossRef DOI)

Publications

Full-text archive of all publications resulting from USGS-funded research available in Publications Warehouse

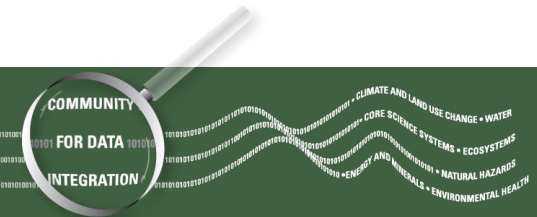
XML Publishing

All USGS series publications available in machine-readable **XML format**.

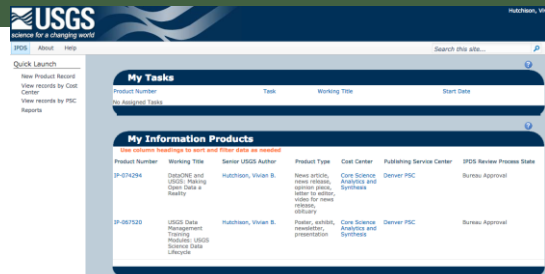


Some accomplishments

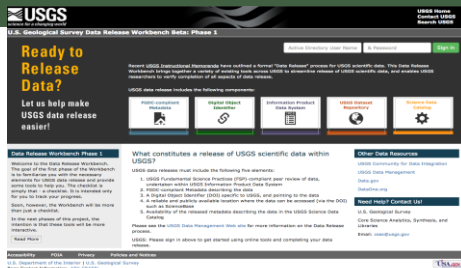
- **Communication:** ~75 presentations reaching ~1,600 employees in 2 months on policies and implementation options
- **Data Policies:** Instructional Memos updated and published as policy; FAQ guidance
- **Repositories:** designed and implemented a pilot for establishing trusted digital repositories in USGS, based on the World Data Center.
- **ORCiDs:** established formal agreement with ORCiD; began process to update systems to accommodate ORCiDs, including Active Directory (still ongoing)
- **CHORUS:** The USGS Joined CHORUS to help monitor for scholarly publications becoming public after an embargo period of one year (if applicable).
- **IPDS:** updates to accommodate a dark archive for publications and data releases
- **Publications Warehouse:** catalog records of all series pubs and external publications; updated data model for ORCiDs, etc.
- **Science Publishing Network:** Completed contract for XML publishing progress
- **Data release applications:** Increased connections between established systems such as ScienceBase data repository and Pubs Warehouse/Science Data Catalog
- **IGSN issuance:** In progress



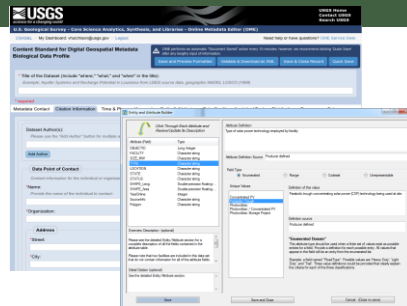
Connecting tools to meet policy and Public Access Plan



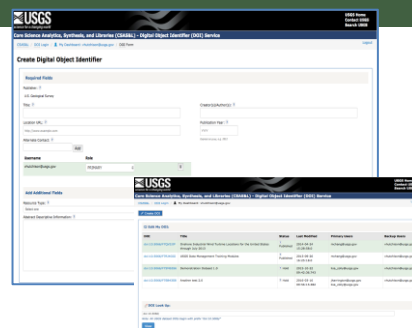
USGS My Tasks and My Information Products interface. The 'My Tasks' section shows a table with columns for Product Number, Working Title, Task, Working Title, and Start Date. The 'My Information Products' section shows a table with columns for Product Number, Working Title, Product Type, Data Center, Publishing Service Center, and SPDS Service Process State.



USGS Ready to Release Data? interface. It features a 'Ready to Release Data?' heading and a 'Let us help make USGS data release easier!' message. Below this, there are several icons representing different data release processes: 'Data Release Process', 'Data Release Process', 'Data Release Process', 'Data Release Process', and 'Data Release Process'.



USGS Metadata Contact interface. It shows a 'Metadata Contact' form with fields for 'Name', 'Email', 'Phone', and 'Address'. There is also a 'Data Point of Contact' section with a 'Name' field and a 'Data Point of Contact' section with a 'Name' field.



USGS Create Digital Object Identifier interface. It shows a 'Create Digital Object Identifier' form with fields for 'Name', 'Email', 'Phone', and 'Address'. There is also a 'Data Point of Contact' section with a 'Name' field and a 'Data Point of Contact' section with a 'Name' field.

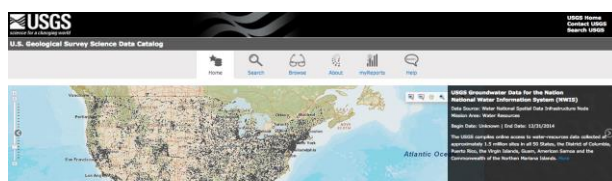
ORCID
Connecting Research
and Researchers

CHORUS

xml



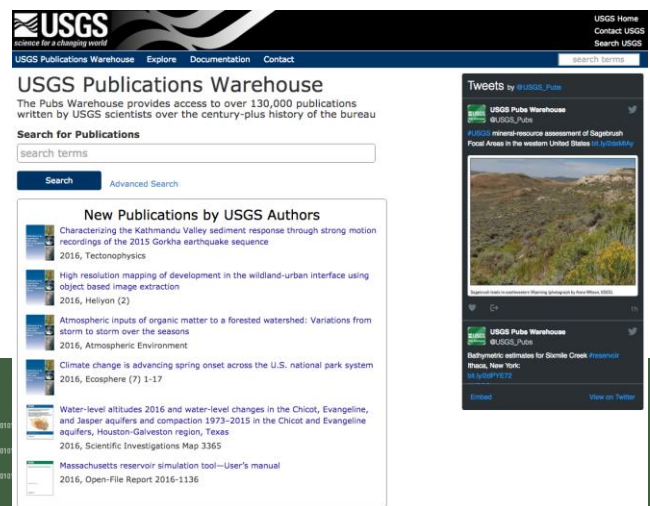
USGS ScienceBase Catalog interface. It shows a search bar with the text 'Type some text to search...' and a 'Search' button. There are also links for 'Log in' and 'Sign up'.



USGS U.S. Geological Survey Science Data Catalog interface. It shows a map of the United States with a search bar and a 'Search' button. There are also links for 'Log in' and 'Sign up'.



USGS Your source for open data within USGS interface. It features a 'Your source for open data within USGS' heading and a 'Frequently Accessed USGS Data' section. The 'Frequently Accessed USGS Data' section includes links to 'The National Map Data', 'Earthquake Information', 'Biodiversity Information', 'South Florida Hydrology', 'Energy Resources Program', 'Water Resources of the United States', 'Ocean Biogeography Information System USA', and 'Palawan Wildlife Research Center'.



USGS Publications Warehouse interface. It shows a search bar with the text 'search terms' and a 'Search' button. There are also links for 'Log in' and 'Sign up'. The interface includes a 'New Publications by USGS Authors' section and a 'Tweets by @USGS_Pubs' section.

Challenges that remain

There are still some challenges with implementing public access, but we're getting closer:

Challenge 1: Ensuring our systems and processes are ready for meeting data release/publication access requirements, including a wide variety of data types (open, sensitive, etc.)

Challenge 2: Assuring proper long-term resourcing of our trusted digital repositories, and consideration for the role of non-USGS repositories

Challenge 3: Large datasets – where to put them, how to manage them for the long term

Challenge 4: Positioning Centers and Programs to be successful in data release/management through training and continued improvements



Essentials for success

- **A shared vision/goal:** everyone understands where we are trying to go
- **Leadership:** Executive champions, accompanied by action-oriented leaders to guide the process
- **Contribution:** Open avenues for individual voices/ contributions create critical buy-in
- **Accountability:** Helpful for ensuring progress happens quickly
- **Communication:** transparency and openness about progress encourages interest and participation
- **Education and training:** Provide support in the trenches to facilitate use and implementation.



Thank you!

Lindsay Powers

National Geological and Geophysical Data Preservation Program

lpowers@usgs.gov

Thanks to Viv Hutchison for the contribution of several slides.