NOAA Persistent Identifiers

Nancy Ritchey
Chief, Archive Branch, Data Stewardship Division
NOAA/NESDIS National Centers for Environmental Information (NCEI)

Thursday, 23 April 2020
WGISS-49
Virtual

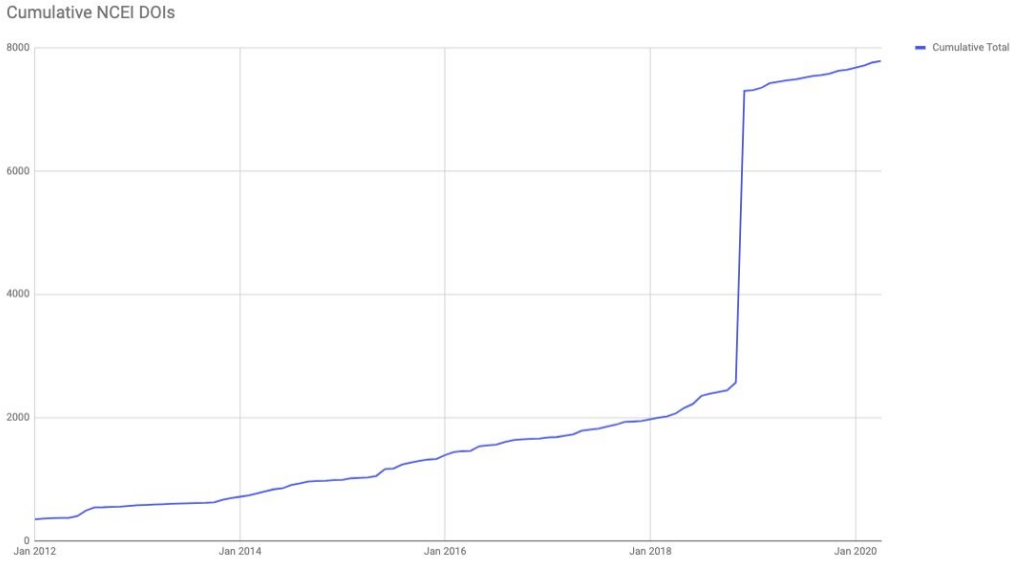
Persistent Identifier Guidance and Governance

- NOAA Data and Publication Citation Procedural Directive (PD) defines basis for Persistent Identifiers (Digital Object Identifiers) for a data collections
- PD assigns responsibilities for Working Group oversight and role of data stewards (data) and librarians (publications) for creating and managing DOIs
- Working Group adjudicates questions and recommends improvements to implementation processes

Data PIDs By the Numbers

Track and report metrics

- ~7800 **Data** DOIs since NOAA started minting Data DOIs in 2012
- Received 1074 transferred DOIs when CDIAC closed
- Average* of ~310 DOIs/year

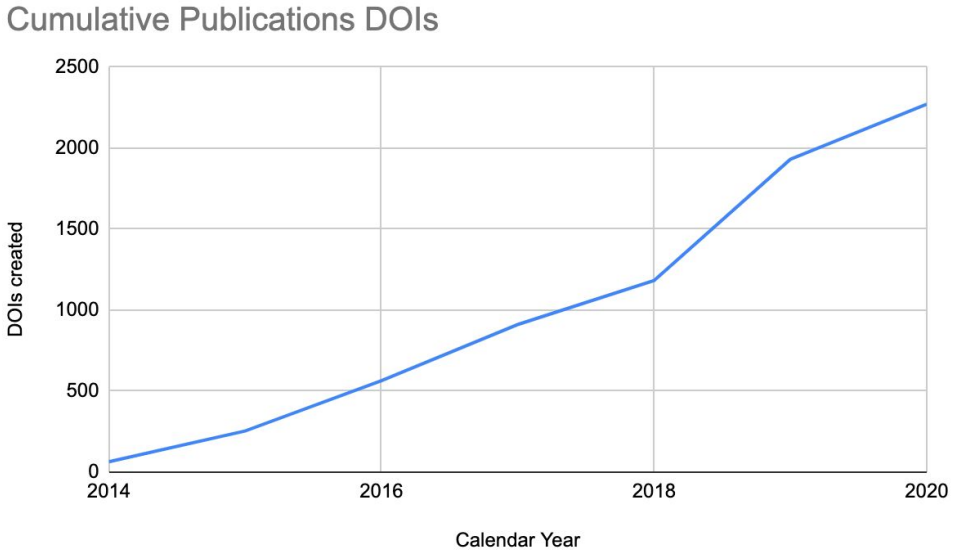


*Excludes ~5000 DOIs in 2018 for special project to support WDS Paleoclimate

Publication PIDs By the Numbers

Track and report metrics

- >2270 **Publication** DOIs since NOAA started minting Publication DOIs in 2014
- Average of ~324 Publication DOIs/year



PIDs for Documentation and Software

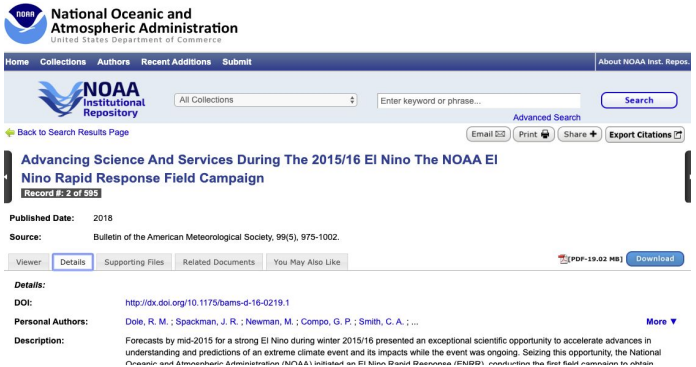
- NOAA Central Library responsible for publications, including NOAA Tech Memos, Cruise Reports, other NOAA publications
- NCEI responsible for data and documentation that may augment or accompany data in archival data collections
- NOAA does not have a centrally-managed software repository and does not currently create a DOI for NOAA-created software

Landing Page Approach

- Data DOIs link directly to a uniformly-presented landing page presenting the content of ISO 19115 metadata for the data collection
- Publication DOIs link directly to the NOAA Institutional Repository landing page for the publication
- Metadata provides two-way cross-references between DATA and PUBLICATIONS when those linkages are known and documented

DATA

PUBLICATION



Recent Developments

- New Version 2.0 of the NOAA Data and Information Citation PD (draft is in review/approval process)
  - Clearer delineation between requirements and processes for NOAA data and publications
  - Documents processes for reserving a DOI for data and publications
- NCEI Data DOI Minting SOP for consistent implementation
- Published basic DOI information on the NCEI archive page
- Planned work: Integrate DOI management into enterprise metadata tools

Lessons Learned

- Important to have a publicly-accessible policy statement (e.g., PD) to underpin the process and basic requirements
- Review and maintain the policy statement to keep up with and manage expectations
- Be prepared (via a team or Working Group) to field and resolve nuanced questions about the process and policy interpretation
- Communicate updates and changes widely and quickly

Other Notable Points

- NOAA changed DOI vendors which affected the DOI prefixes used for NOAA data and publications and the process for managing and reporting DOI metrics
- Management and monitoring of DOI minting across NOAA due to license limits (10K/year)
- Challenges for managing metadata for data that have a pre-existing DOI from a non-NOAA source

Questions/Challenges

- Event-based DOIs for cruises have been requested, but not adjudicated. Do other repositories mint DOIs for events?
- How should data citation information be embedded in distributed data files? Like a watermark for data.
- Users access subsets of various data. How can dynamic, custom citation information be provided to users for their ad hoc data access?
- Support of DOIs required by publishers in order to publish articles. How are you handling that?