EPA Data Governance Council

Meeting Summary – November 16, 2020

Revising the Enterprise Information Management Policy (EIMP)

- EPA's Chief Data Officer (CDO) provided an update on the Agency's effort to revise the Enterprise Information Management Policy (EIMP).
- The EIMP is the guiding policy on information management and metadata publishing. The overarching policies will have guidance documents and procedures attached.
- In the next couple of months, EPA's Office of Information Management (OIM) will distribute the draft, revised policy to the Data Governance Council (DGC) for review and feedback.

Data Skills Assessment

- The Data Skills Assessment will be piloted with Region 1, beginning on January 20, 2021.
- The subcommittee is developing the questions the analysis should answer.
- The CDO briefed EPA's Chief Operating Officer (COO) to ensure awareness of the effort at the Agency leadership level.

Data Webpage Update

- The primary goals of the EPA Data Webpage update effort are to streamline data communication with the public and create a narrative that highlights data usage.
- The workgroup is beginning with external-facing pages, followed by the intranet pages.
- The DGC reviewed draft wireframes that highlight the proposed format for the revised data webpages.

Metadata Management COTS Analysis

- COTS analysis group held a kickoff meeting on November 9, 2020. The meeting included the following:
  - A presentation from EPA's CDO on EPA's data landscape, challenges and the group's charge.
  - A presentation from EPA's OIM on data catalog challenges and features.
  - A presentation on the Office of Research and Development's (ORD) ScienceHub catalog.
  - A discussion of the proposed evaluation plan and schedule.
- Next steps include defining the scope of analysis, evaluation strategy, requirements, and evaluation criteria.
- DGC members discussed that the effort is focused on what the data is and how it is structured. Due to mergers and development of the platforms, many of tools include components that deal with interoperability or attaching to controlled vocabularies.
- DGC members also discussed relationships between this effort and the Agency's implementation of a new records management system. The new records management system should be addressed as much as possible. The ideal result would be a data management metadata cataloguing system that fits as many parameters across the agency as possible.
- DGC members discussed the benefits of looking at what solutions have been used at other federal agencies. A suggestion was made that USGS has a good system: https://www.usgs.gov/products/data-and-tools/data-management/data-release

- A suggestion was made to look at machine learning (ML) and other automation solutions that make it easier to use when evaluating options. It was noted that a lot of these products have artificial intelligence (AI) built in to help organize data and reduce duplication.

Data Inventory Subcommittee/Open Data Plans

- The charge of the subcommittee is to ensure EPA's data inventory is as complete as possible.
- The subcommittee will need to define what needs to be in the inventory. EPA knows the Environmental Dataset Gateway (EDG) is currently incomplete.
- One of the early tasks is to define common definitions.

Environmental Dataset Gateway (EDG) Overview

- The DGC was provided with an overview of EDG, EPA's metadata catalog for data sets.
- The purpose of the EDG is to provide EPA staff with a single access point to EPA's geospatial and non-geospatial recourse and metadata records. It is also meant to meet legislative and regulatory mandates and agency requirements for metadata compliance, sharing and management.

Action Items

- The CDO will reach out to the Office of Chemical Safety and Pollution Prevention (OCSPP)and ORD to set up data discussions.