

EPA Data Governance Council

Meeting Summary – January 31, 2022

General Updates

- The metadata management tool was selected. EPA is now going through the FITARA process to procure the selected tool. The goal is to deploy the tool in Amazon Web Services (AWS) and begin migrating existing EDG metadata within the next few months.
- The Office of the Administrator and the Office of the Chief Financial Officer are developing a plan to meet reporting requirements under the new Infrastructure Law.
- The Agency Equity Team sent their action plan to OMB in January.
- Richard Allen will share the finalized group of Priority Questions with the DGC.
- The IT/IM conference next month will include a data session focused on grants.
- The Enterprise Data Management Policy (EDMP) will be shared for Agency-wide review next month.
- Ongoing DGC project updates are available in the 01/31/2022 DGC Project Updates wiki.

Data Governance

- Heather Barnhart discussed the near-term operational plan due to OMB on a regular basis.
- The Data Governance Subcommittee is working on a streamlined governance and inputting identified activities into the near-term operational plan.
- The first few years of the operational plan focus on foundational activities such as governance, planning and infrastructure.
- The Subcommittee based planned activities on findings from the data skills survey, the data maturity assessment and the DGC's Strengths, Weaknesses, Opportunities, Threats (SWOT) discussions.
- The FY22/23 Operational Plan is organized around strategic goals, including open data, culture, infrastructure and partnership.
- The Subcommittee is meeting next week to finalize the operational plan. They will then gather feedback from the Data Maturity Subcommittee and the DGC. The goal is to finalize the plan by early March.

Data Maturity Assessment

- Richard Allen reviewed high level takeaways from the Data Maturity Assessment.
- EPA used Gartner's IT Score for Data and Analytics to conduct the Data Maturity Assessment.
- There were 100 respondents from 11 EPA organizations.
- EPA's overall maturity score is 2+ on Gartner's IT Score scale.

Proposed DGC Subcommittees for 2022

- The CDO requested engagement in DGC subcommittees for 2022 to assist in meeting OMB required actions.
- DGC members should complete and share the DGC Subcommittee Interest Survey. Subcommittee members do not need to be members of the DGC.
- The following existing DGC Subcommittee will continue in 2022:
 - Data Inventory

- The DGC is required to Publish and Update Data Inventories.
 - The Subcommittee will ensure that the enterprise data inventory is as complete as possible by adding missing priority data assets to the Environmental Dataset Gateway.
 - Data Governance
 - The DGC is required to develop an enterprise data strategy, data governance plan, and publish EPA's Open Data Plan.
 - Data Skills Training
 - The DGC is required to identify opportunities to increase staff data skills. The Subcommittee will build upon 2021 survey findings to review existing and potential future training courses to define a curriculum based on skill level and job responsibilities to strengthen data skill attainment.
 - Data Maturity
 - The DGC is required to assess data and related infrastructure maturity.
 - The Subcommittee will analyze 2021 assessment results to set DGC priorities for CY 2022-2023 and use outcomes in FY 2023 strategic planning.
- The following new DGC Subcommittees are proposed for 2022:
 - Artificial Intelligence (AI)
 - The AI subcommittee will address data challenges specific to AI. Examples include using AI to better understand and profile data, supporting methods to optimize data for AI, efficient large data processing methods, integration of disparate datasets, categorization of data, and usage of unstructured data.
 - Related topics will include ethical use of AI and support for an AI ready workforce.
 - Communications
 - Develop communication plans to identify the engagement and key messages to various stakeholders around EPA's Data Strategy and Data Governance
 - Create baseline communication materials to support the DGC members in their roles to engage stakeholders within their organization
 - Metadata Stewardship
 - The Subcommittee will bring together and foster a vibrant, active network of metadata stewards that work with the Data Inventory Subcommittee to define how to maintain comprehensive and current metadata for priority data assets.
 - Data Sharing CRADA
 - Steven Fine presented an overview of the Data Sharing CRADA subcommittee.
 - EPA is working on potential agreements with Amazon and Microsoft for cloud sharing of large datasets. These providers have open data programs, but their standard agreements are not acceptable for the Agency.
 - The goal is to establish a cooperative research and development agreement with Amazon and Microsoft because they are the largest providers and EPA has existing relationships with them.