EPA Data Governance Council

Meeting Summary – January 29, 2024

General Chief Data Officer Updates

- Artificial Intelligence Subcommittee Chair, Andrew Chase updated the Data Governance Council on the Large Language Model pilot being planned to test the use of OpenAI within EPA.
    - The goal of the LLM pilot is to use the Microsoft system GovChat within an EPA framework.
    - How participants will be selected for the study has not been determined.
- Chief Data Officer, Richard Allen, updated the Council on the progress of establishing a system of governance to review, approve and inventory AI systems implemented within EPA.
    - The Federal Information Technology Acquisition Reform Act process has been updated to include AI checkmarks.
    - The Office of Research and Development Scientific Onboarding Process may also be updated to include AI checkmarks.
    - There will be another data call conducted in the spring of 2024 to ensure the required AI inventory is reported accurately.
- EPA has a signed Memorandum of Understanding for the Amazon Web Services Open Data Program. The next step in the implementation process will be to determine how data is loaded and submitted to the program. Anyone interested in hosting data sets in the AWS system should contact their Data Officer.
- The Data.gov metadata standard is being updated.

Environmental Justice Data Standards and Cumulative Impacts

- Patricia Koman gave a presentation on Environmental Justice Data Standards and Cumulative Impacts.
- Cumulative Impacts recognizes that communities can experience multiple consecutive harms that make them more vulnerable to climate change.
- Multiple drivers led to the creation of an agency-wide Cumulative Impacts Workgroup which is setting up workplans for projects across the agency.
    - OAR is implementing a three-part CI strategy.
    - The workgroup has a callout for contractor support to advance CI methods in regions and offices.

Integrated Compliance Information System Subcommittee

- Jane Wallace presented a request to form an Integrated Compliance Informational System Modernization Subcommittee under the DGC.
- In addition to the subcommittee, a Fusion Team will be created as an integrated project team that includes expert partners from states, Tribes, the EPA and counterparts in IT implementation.

- The Office of Enforcement and Compliance Assurance's ICIS is EPA's principal compliance and civil enforcement data asset.
- ICIS data and data models are generated by EPA, state and local sources, and regulated entities.
- The ICIS modernization goals include:
    - Meeting the business needs of users.
    - Developing a user centered design to make reporting easier for users.
    - Improving the base technology of the system.
    - Meeting compliance assurance reporting requirements.
    - Reducing costs and time for operations, maintenance and future developments.
    - Generating better data products.
- The ICIS Modernization Workplan includes:
    - Gathering system requirement information from high level users.
    - Analysis of alternative systems.
    - Designing and developing the shared elements of the CORE ICIS system.
    - Delivering and iterating CORE ICIS modules.
    - Designing and delivering non-CORE ICIS modules.
- The DGC ICIS Subcommittee will assist in the modernization plan by developing use cases, exchanging data, and developing data and functional models.

Internal Data Site Updates

- Connor Griffin announced that the Communication Subcommittee completed revisions to the public epa.gov/data site and presented updates for the Internal Data Site project.
- The internal data site will have an Intranet component for static information and a SharePoint/Teams component for connecting communities and allowing dynamic communications.
- Site content will include:
    - Data Stewart Quick Start Guide.
    - Data Skills Training Information.
    - Data Tools.
    - Communities of Practice Information.
    - Data Success Stories.
    - Data Standards.
    - Data and Information Architecture.
- Connor Griffin asked DGC members to contact him with suggestions for any additional data related information, links or tools that should be included in the new internal site.

Data Skills Subcommittee Updates

- Data Skills Subcommittee Lead, Janet Kremer, presented updates from the Data Skills Subcommittee.
- The Early Learner/Beginner Persona Training page was launched in August 2023.

- In February 2024 the subcommittee will launch a new Data Skills Landing Page and the Data Apprentice Persona Training page.
- Data Journeyperson and Data Expert Persona Training pages are both scheduled for launch later in 2024.
- Janet provided a preview of the new pages that are being prepared for launch.
- Each page will have a feedback survey available so users can indicate whether the offered trainings are fulfilling their needs.

Enterprise Data Catalog Updates

- Janet Kremer presented updates about the EDC.
- The EDC is replacing the Environmental Dataset Gateway which was decommissioned in 2023.
- Previously the system was called Alation which is the name of the metadata management tool the system uses.
- The EDC will be deployed in March and include the current assets in the Enterprise Data Inventory.
- The launch will include a demonstration of capabilities that allow users a preview of the database format and structure.
- Everyone with EPA credentials will have a Viewer Role in the EDC system but the licenses required to edit information within the EDC will be limited to Data Officers and Data Stewards. Additional access roles will be assigned based on need.
- There will be a data call completed in February asking data owners to update their Enterprise Data Inventory information.
- The EDC team will train Data Officers to update, manage and use EDC data for their office and region.
- The EDC team will provide EPA staff training for viewing and using EDC data through EPA developed training resources, Alation training resources and direct training by request.

Data Officer Priorities

- Chief Data Officer, Richard Allen, presented updates on the Data Officer status and priorities.
- Data Officers were identified in August 2023 and held their first meeting in September 2023 which focused on the background and context of the Data Officer position.
- Subsequent meetings have been held to gather information from the DOs and create a list of refined priorities.
- Moving forward, the DOs will be finalizing the prioritized tasks and updating the Enterprise Data Catalog for the launch of the EDC.
- The CDO team is working to develop marketing materials to communicate about current DGC projects and the roles and responsibilities of DOs and Data Stewards.