EPA Data Governance Council

Meeting Summary – May 24, 2021

Chief Data Officer (CDO) Updates

- Review updates on ongoing activities on the DGC Updates tab on the Data Governance Council Team site.
- The Metadata Management COTS Analysis Workgroup selected three vendors to further evaluate as part of an additional evaluation phase. The goal is to select a software by August in order to begin implementation in Quarter 1 of Fiscal Year 2022.
- The Data Inventory Subcommittee is relaunching with goals to ensure EPA's data inventory is more complete and to build up the Agency's community of data stewards.
- The process of updating the Enterprise Information Management Policy (EIMP) to reflect the Open Government Data Act will begin this summer. The project manager is looking for volunteers to participate on a stakeholder group to review and comment on the draft revisions.
- The DGC is starting to build an Agencywide governance plan and will kick off the Data Governance Subcommittee in the next few months.
- The IT Portfolio Review data discussions are ongoing. The focus of the data discussions is on existing data management practices across the Agency.
- The Learning Agenda is moving forward and aligned with the strategic plan.
- The Equity Data Workgroup kicked off in April 2021. The Data Subgroup will be requesting information from the DGC in the next few weeks.
- The Data Skills Assessment Survey was released on April 19, 2021 and closes on May 24, 2021. There is a good distribution across program offices and regions in the 1,647 responses received.

Data Success Story: Data Science Training Program

- Carrie Cummings and Daniel Ginsberg provided an overview of their experience participating in OMB's Data Science Training Program.
- OMB's Data Science Program aims to upskill the federal workforce, train to apply data science techniques to enhance data gathering, analysis and presentation with data visualization to make informed, data-driven decisions.
- As a part of the program, they conducted a capstone project utilizing Machine Learning to identify, predict and visualize patterns related to an emerging contaminant-of-concern, Perfluorooctanoic acids (PFAS), in well and groundwater samples.
- The group noted that utilizing cross-functional teams is key to good data analysis and visualization. A combination of subject matter expertise, statistical knowledge and computing skills is needed for success.