

Data Governance Transformation: Now is the time to revisit Data Governance Policy

Jacob M. Pasner

Assistant Research Professor - Georgetown Massive Data Institute

This talk presents MDI's Data Governance Transformation (DGT) project, an already-in-use data governance policy & infrastructure framework developed to support robust implementation of Privacy Enhancing Technologies (PETs) in response to precedent shattering data use in the United States. As government agencies increasingly rely on complex and distributed data ecosystems, traditional data management approaches have proven insufficient to ensure data quality, accessibility, privacy, and interoperability. This outdated data governance model makes it difficult for staff and consultants to support robust implementation of privacy preserving technologies due to the lack of transparency on requirements from siloed program staff and inconsistent statutory interpretation. This in turn leads to slowed development requiring multiple change requests as each new PET project is forced to “re-invent the wheel”. The DGT project proposes to cut short this inefficient, frustrating, and costly cycle by establishing standard policies which ensure mission critical data governance information is prioritized. As a result DGT aims to support more rapid and robust adoption of PETs in addition to meeting the growing public demand for accountability and transparency in the age of Artificial Intelligence and politically motivated shifts in government data usage norms.

In order to accomplish this the DGT framework establishes a number of structures. First we establish a Data Governance Chain of Command which integrates with existing organizational structures to oversee the roles and responsibilities of Data Governance Teams, their interactions with IT staff and leadership, and General Counsel who support standardized statutory interpretation. Next we define a reference data architecture which establishes infrastructure naming conventions to ensure that inter-agency development conversations use common language to avoid confusion. Finally a RACI (Responsible, Accountable, Consulted, Informed) matrix goes into the details of connecting the defined chain of command roles with their respective responsibilities within the reference architecture. To bring all of this together we propose the use of a data catalog which centralizes all data information integrated with Role Based Access Controls to ensure only authorized staff can view sensitive or private information. Taken together we propose that these structures provide the necessary framework for all Data Governance activities spanning day-to-day operations, data sharing using PETs, transparency and oversight requirements, and the development of trustworthy AI.

Next we go on to provide practical “Getting Started” guidance to help agencies assess their current data governance system and identify early steps to build momentum for a full data governance transformation. This guidance includes developing artifacts that become critical to data projects such as conducting a data inventory which includes privacy/sensitivity categorization and connecting each data element to its statutory requirements on privacy, security, and permissible uses. Note that this baseline information is critical to real world implementation of PETs but is often lacking which leads to implementations which fail to meet business requirements. Ultimately this

provides a roadmap for agencies to establish and mature their metadata curation pipeline, cloud infrastructure conventions, privacy requirement handling, and related contracting requirements such as data portability and code version control. In short, DGT aims to help agencies re-establish control over the governance of their own data thus decreasing their dependency on contractors for data governance.

In conclusion, we propose this framework for discussion and eventual adoption as a replicable model for both government and non-government entities seeking to implement or refresh their data governance strategies to meet the demands of modern public service delivery including robust privacy, transparency, and accountability.

References:

1. [Data Governance Transformation Project at Georgetown's Massive Data Institute](#)
2. [DGT Policy Document](#)
3. [DGT Chain of Command](#)
4. [DGT Reference Architecture Diagram](#)