

DATA GOVERNANCE RESOURCE INDEX



This is a data governance reference document that contains links and citations; abstracts for websites, articles, white papers and guides; frequently asked questions; and a glossary.



Contents

Introduction	2
Websites and Professional Organizations.....	3
White papers and Guides.....	5
Data Governance in Other States	7
Guides and Case Studies	7
Websites	8
Frequently Asked Questions	10
Glossary of Common Terms.....	13

This publication is issued by the Office of Management and Enterprise Services as authorized by Title 62, Section 34. Copies have not been printed but are available through the agency website. This work is licensed under a Creative Attribution-NonCommercial-NoDerivs 3.0 Unported License.

Introduction

The need for and usefulness of data governance is growing across all industries, including local, state and federal governments. As Oklahoma state agencies work to design and implement data governance programs and dedicate staff to data governance and data management efforts, it is important that those agencies are provided the support they need to be successful. Modeling programs on successful structures and incorporating best practices established across organizations and industries will help ensure data governance programs will last and tax dollars are well spent.

This document includes a variety of relevant resources from government and nongovernment organizations spanning all sizes and myriad industries. It functions as a starting place in your search for resources and networking opportunities with other data governance professionals. The goal of this document is to help you learn about industry standards and best practices so that they may be built into the foundations of successful data governance programs across Oklahoma state agencies.

With this index, the first step in building a successful and long lasting data governance program is as easy as clicking a link or downloading a document. Let the resources provided here serve as your starting place for learning all you can about data governance.

Websites and Professional Organizations

Developing and implementing a data governance program can't be done in a silo. Luckily, there are a number of professional organizations that create a community of data governance professionals ready to share their in-depth knowledge, vast experience and lessons learned as you stand up your agencies' data governance program. The websites and professional organizations listed below offer support with online resources, informational videos, best practices and networking and educational opportunities all focused on data governance and related topics.

Data Management Association International | <https://www.dama.org/>

DAMA International is a professional organization providing research, education, publications, standards' promotion and activities to enhance the practice of data management.

Data.gov | <https://www.data.gov/>

Data.gov provides data, tools and resources to conduct research, develop web and mobile applications, design data visualizations and more.

Data Governance Insider | <http://data-governance.blogspot.com/>

Data Governance Insider is a blog containing articles and references on a variety of topics relating to data governance, data management and related disciplines.

The Data Governance Institute | <http://www.datagovernance.com/>

DGI is a source of in-depth, vendor-neutral data governance best practices and guidance including The DGI Data Governance Framework.

Data Governance on YouTube | https://www.youtube.com/results?search_query=data+governance

Searching for data governance on YouTube returns a variety of informational videos.

Data Governance Professionals Organization | <http://www.dgpo.org/>

DGPO is a nonprofit, vendor-neutral association of business, information technology and data professionals dedicated to advancing the discipline of data governance.

DATAVERSITY | <http://www.dataversity.net/>

DATAVERSITY provides high quality educational resources for business and IT professionals on the uses and management of data.

Government Technology | <http://www.govtech.com/>

Government Technology covers IT's role in state and local governments. Coverage includes IT case studies, emerging technologies and the implications of digital technology on the policies and management of public sector organizations.

GovLab | <http://thegovlab.org/>

A search on **GovLab** will return articles and blog posts on data governance topics from a government perspective. GovLab's goal is to strengthen the ability of institutions, governments and people to work more openly, collaboratively, effectively and legitimately to make better decisions and solve public problems.

National Association of State Chief Information Officers | <http://www.nascio.org/>

NASCIO provides members with products and services designed to support the challenging role of the state CIO, stimulate the exchange of information and promote the adoption of IT best practices and innovations.

White Papers and Guides

Designing a data governance program doesn't mean developing innovative structures and processes from scratch. Many data governance and data management best practices already exist. The white papers and guides listed below provide in-depth information and best practices on a variety of topics, including data governance's importance, effective strategies for implementing a program, data governance frameworks and data stewardship.

The DAMA Guide to the Data Management Body of Knowledge

The Data Management Association's guide is a compilation of principals and best practices. It provides data management and IT professionals, executives, knowledge workers, educators and researchers with a framework to manage their data and mature their information infrastructure.

The DAMA Guide to the Data Management Body of Knowledge, Edited by M. Brackett, S. Early and M. Mosley. Bradley Beach, NJ: Technics Publications LLC, 2009 (Available for purchase at <https://www.dama.org>).

Seven Steps to Effective Data Governance for State and Local Government Agencies

This white paper explores why data governance is so important in the government sector, and provides seven steps for approaching data governance in your organization. It will also provide real world examples of how data governance was successfully applied in state and local government agencies.

Seven Steps to Effective Data Governance for State and Local Government Agencies. Information Builders. Last modified 2011.
http://www.informationbuilders.com/about_us/whitepapers/download_form/4558.

Data Quality 101: The Ultimate Guide for Data Stewards

This guide offers insight on building data stewardship initiatives in order to improve the reusability, accessibility, and quality of data.

King, Elliot. *Data Quality 101: The Ultimate Guide for Data Stewards*. Melissa Data Corp., 2013.
http://www.melissadata.com/white_paper/data-quality-stewards-ebook.asp.

Data Governance – Managing Information as an Enterprise Asset

This introductory brief presents data governance as a necessary enterprise initiative that must be adopted within the culture of state government. It challenges state government to place data governance as a priority to be investigated and addressed.

National Association of State Chief Information Officers. *Data governance – Managing Information as an Enterprise Asset, Part I: An introduction*. 2008.

<http://www.nascio.org/Portals/0/Publications/Documents/NASCIO-DataGovernance-Part1.pdf>.

The ROI of Data Governance: Seven Ways Your Data Governance Program Can Help You Save Money

This white paper outlines seven ways data governance and stewardship programs can help manage costs and a mechanism for quantifying the return on investment for those contributions.

Thomas, G. *The ROI of Data Governance: Seven Ways Your Data Governance Program Can Help You Save Money*. 2009. http://www.cio.com.au/white_paper/323775/the-roi-of-data-governance-seven-ways-your-data-governance-program-can-help-you-save-money/.

How to Use The DGI Data Governance Framework to Configure Your Program

This white paper offers data governance professionals a visual reference and guidance on using The DGI Data Governance Framework to configure their data governance program.

Thomas, G. *How to Use The DGI Data Governance Framework to Configure Your Program*. 2014. http://www.datagovernance.com/wpcontent/uploads/2014/11/wp_how_to_use_the_dgi_data_governance_framework.pdf.

Data Governance in Other States

State governments nationwide are realizing data are an enterprise asset, much the same as desks, chairs, computers and any other tangible item labeled “state property.” Although a paper label can’t be affixed to data, data can and should be managed and utilized so that it doesn’t sit in a warehouse collecting virtual dust. The guides, case studies, and websites listed below provide real-world examples of what other states are doing in the areas of data management and data governance.

Guides and Case Studies

Geospatial Data Governance Plan GIS Project

This plan outlines a process for geospatial data sets Colorado and describes how this process should integrate and align with the enterprise data governance.

State Geospatial Data Governance Work Group, C. Brown, J. Gottsegen, W. Johnson, M. Rigiroszi, B. Sacco and M. Sullivan. *Geospatial Data Governance Plan GIS Project*. 2010.
http://hermes.cde.state.co.us/drupal/islandora/object/co%3A12946/datastream/OBJ/download/Geospatial_data_governance_plan_GIS_project.pdf.

Maricopa Integrated Health System Uses Data Governance to Align the Enterprise, Increase Productivity and Deliver On Its Mission

This case study details how Maricopa Integrated Health System, in south-central Arizona, uses data governance to align the enterprise and increase productivity.

McKesson. *Maricopa Integrated Health System Uses Data Governance to Align the Enterprise, Increase Productivity and Deliver On Its Mission*. 2015.
<http://www.mckesson.com/healthcare-analytics/resources/case-studies/learn-how-to-use-data-governance-to-align-the-enterprise-and-increase-productivity/>.

Effective Cross-Jurisdictional Collaboration - Governance is Critical!

Nebraska’s education system is used as an example of effective governance, having all public and private K-12 and higher education entities interconnected.

National Association of State Chief Information Officers. (2013). *Effective Cross-jurisdictional Collaboration - Governance is Critical!* 2013.
<http://www.nascio.org/Portals/0/Publications/Documents/EffectiveCrossJurisdictional.pdf>.

Michigan Department of Transportation Safety Data Processes and Governance Practices

This case study was conducted by the National Cooperative Highway Research program to summarize how the MDOT incorporates a data governance structure into their standard business operations and how it has helped them to improve their safety data systems and processes.

U.S. Department of Transportation. *Michigan Department of Transportation Safety Data Processes and Governance Practices*. 2015. FHWA-SA-15-059.

http://safety.fhwa.dot.gov/rsdp/downloads/miDOT_casestudy_dm_final.pdf.

Websites

Colorado Office of Information Technology | <http://www.oit.state.co.us/cto/cim/data-management>

The Colorado Office of Information Technology has a data management program with the purpose of leveraging data and information as enterprise assets and establishing standards and processes to enable more agile solutions and government services.

Georgia Technology Authority | <http://gta.georgia.gov/enterprise-governance-and-planning-main-page>

Georgia Technology Authority's Enterprise Governance and Planning Division promotes an enterprise approach to technology by establishing statewide policies, standards and guidelines based on industry best practices and federal requirements.

Iowa Department of Public Health | <http://idph.iowa.gov/publichealthdata>

The Iowa Department of Public Health has a data management program that includes internal data governance as well as a public health tracking portal, data requests and informatics.

Kentucky Department of Education | <http://education.ky.gov/districts/tech/Pages/KDE-Data-Governance.aspx>

The Kentucky Department of Education has a cross-agency coordinated data governance structure designed to promote data quality, appropriate data use and data security.

State of New Jersey | http://www.state.nj.us/it/ps/Shared_IT_Architecture.pdf

New Jersey has a data governance program focused on defining and classifying data objects and assisting in resolution of data quality issues, so that state government can be more efficient.

Texas Health Human Services Commission | http://www.hhsc.state.tx.us/hhsc_projects/oehc/index.shtml

The Texas Health and Human Services Commission has an enterprise data program that provides an organization and policy framework to support data management standards across the state's health and human services agencies to improve data sharing and data quality.

Virginia Information Technologies Agency | <https://www.vita.virginia.gov/oversight/dm/>

The enterprise information architecture of VITA promotes the governance, asset management and sharing of Virginia's data assets.

Wyoming Department of Enterprise Technology Services | <http://ets.wyo.gov/governance>

The Wyoming Department of Enterprise Technology Service has an IT and data governance program that identifies the decision rights and accountability framework to encourage desirable behavior in the use of IT.

Frequently Asked Questions

1. **How is data governance different from project and portfolio governance and IT governance?**

Data governance complements project and portfolio governance and exists under the overarching structure of IT governance. Project and portfolio governance focuses on defining a portfolio of investments, setting performance objectives and evaluating and managing risk for IT projects. Data governance, on the other hand, focuses on creating a structure that will enable the organization to align data management efforts to business objectives, support regulatory compliance and manage the risks associated with managing data. IT governance is the overarching structure that focuses on the overall IT Infrastructure, including managing resources and risks, to ensure IT efforts and solutions align with the organization's mission and business goals. To borrow an analogy commonly used by the data management community: IT governance focuses on the pipelines in the organization's IT infrastructure, data governance and project and portfolio governance focus on the water that flows through those pipelines.

2. **What is the difference between data governance and data stewardship?**

Data governance is a strategic function. It is strategic in the sense that it is long-term, general and global. Examples of data governance tasks include the creation of a structure for participation (the committees, working groups and councils for the data governance program), defining the goals and principles, establishing a communications plan, defining the policies and processes, and defining the roles and responsibilities.

Stewardship is a tactical function. It is tactical in the sense that it is specific, local and may be short term. Examples of tactical data stewardship tasks include defining the data (identifying key data, gathering definitions, documenting allowable values), defining business rules (for creation of data, for usage of data, for derivation of data), documenting data sources, setting data quality targets, metadata identification and documentation, and remediation of data issues.

3. **How do we decide who the data stewards should be?**

Data stewards should be those who are subject-matter experts in their respective data domains and who consult with and support business unit staff in their day-to-day data management responsibilities.

4. **Should we tie ownership and stewardship to data types?**

Accountabilities may be tied to a type of data that may be:

- Master data.
- Transactional data.
- Reference data.
- Metadata.
- Historical data.
- Temporary data.
- Or other types.

Best Practice: Most organizations answer “yes” to this question. Stewards or others who are assigned data-related responsibilities are expected to work with only one or a few types of data rather than all types.

5. **For what data subject areas will we first assign ownership or stewardship?**

Information-related or metadata-related accountabilities that focus on master data may be tied to different subject areas, such as customers, products, locations, organizational hierarchies, etc.

Best Practice: Data governance pilot projects often strive to govern a manageable set of data elements within a single subject area. Accountabilities are assigned to standardize data elements, specify and enforce valid values, and address data quality.

6. How should we assign ownership or stewardship to data subject areas?

Some organizations assign an enterprise data steward with ultimate accountability for data within a subject area or domain. Others create communities of data stewards and others who work with that data. Another approach is to tie accountabilities to a master data management program rather than to stewardship. And still another approach is to assign data-related responsibilities to functional roles rather than to stewards.

7. At what level of granularity should we assign ownership or stewardship?

- Documents.
- Content units (used in documents, web displays, reports, etc.).
- Data feeds.
- Data records.
- Raw data.
- Domains of data (for example, all data related to customers).
- Usage-related collections of data (for example, all fields appearing on a certain report or all fields included in a compliance mandate such as HIPAA, HMDA or Sarbanes-Oxley).
- Specific data entities (for example, within a data feed, an entire customer record, including the customer's ID, name and all related data).
- Data attributes (for example, only a certain preference flag within a customer record).

Best Practice: Most organizations getting started with data governance and stewardship feel that assigning all levels of granularity simultaneously is a “boil the ocean” type of mistake. Instead, they choose certain levels of accountability for certain data, and then expand scope over time.

8. Should we tie data ownership and stewardship to processes and data flows?

Some organizations assign just one data owner or data steward for a data element or subject area. This person is responsible for the data no matter where it appears in an organization. This approach is not feasible for most organizations, however, with complicated data flows. An alternative is assigning accountabilities for only a few segments in a data flow. One or more data stewards or subject-matter experts could be responsible for access control, quality or typical master data responsibilities for specific data within those segments.

9. Should we tie data ownership or stewardship to compliance or usage?

Some organizations assign accountabilities for related sets of data. For example, HIPAA requires protections of personally identifiable information; some organizations put teams in place to locate that data across systems, to specify controls for the information and to monitor compliance. Likewise, some lending institutions may assign accountabilities to review all data subject to Home Mortgage Disclosure Act compliance.

10. Why does data need to be managed as a state asset?

Data are a valuable state resource. It has real, measureable value. The data's primary purposes are to aid in decision-making and to provide accountability and responsibility.

11. Who will have access to the data, and how will it be used?

Defining data access rules falls within the scope of defining policies, procedures and business

rules. This is when and where to determine who will need access to what data.

12. Why does data include standard metadata?

Common deployment of data documentation schemes promotes data reusability, reliability and the possibility of sharing across the department.

13. What is the relationship between data governance and data quality and do you need both?

Since data governance and data quality rely very much on each other. You would not want to do one without the other if you want to successfully manage and improve the quality of your data in a sustainable manner.

14. Which data entities and data elements should be governed at the enterprise level?

It is impractical to govern every data element, so focus should be on those that are deemed critical for business operations (financial reporting, various external disclosures, risk management, accounting, etc.), decision-making and reporting purposes. It is necessary to engage subject-matter experts within each line-of-business and corporate support function to identify the key business processes and the associated critical data elements. Focus must be on governing this set of enterprise critical data at the enterprise level and to not boil the ocean.

15. What roles do the business, operations and technology organizations play in implementing data governance?

Business organizations have a deeper understanding of data, its definition and usage, decision support, modeling, risk management and reporting, industry standards and alignment, and other such aspects. Business is also aware of the ramifications of data quality issues and inconsistencies in its application to the bottom line. Therefore, as subject-matter experts they should be assigned the role of data trustees and data stewards, with technology and operations teams playing a critical data custodian, trusted adviser and implementation role, to ensure that the right systems, infrastructure and processes are deployed to support and sustain data governance.

16. How should data-related issues be logged and addressed?

Proactive identification and management of data-related issues is required to lower systemic impacts. Each enterprise critical data element should be tagged with its system of record, trusted source, data steward and other pertinent metadata, to facilitate root cause analysis and remediation of issues. Issues should be logged in an enterprise issue management system and assigned to the respective steward, whose role is to triage the issues, drive root cause analysis, assign them to the appropriate owner (data, process or technology) and ensure that they are resolved per agreed upon service level agreements. The governance team should mine issue-related data to find patterns of data anomalies, run predictive analytics on the impact of such issues to downstream systems and provide aging reports to management.

17. How do organizations sustain data governance programs over the long haul?

Implementing the governance model, assigning roles and responsibilities, rolling out corporatewide standards and policies related to data, creating an organizational structure and appropriate escalation mechanisms, proactively monitoring compliance to standards and policies, communicating the value of the program to all stakeholders and continuously improving the process, technology and people aspects of data governance will ensure its sustenance over the long haul.

Glossary of Common Terms

Access Management: A discipline that focuses on ensuring that only approved roles are able to create, read, update or delete (CRUD) data — and only using appropriate and controlled methods. Data governance programs often focus on supporting access management by aligning the requirements and constraints posed by governance, risk management, compliance, security and privacy efforts.

Audit: An independent examination of an effort to determine its compliance with a set of requirements. An audit may be carried out by internal or external groups.

Audit Trail: A record that can be interpreted by auditors to establish that an activity has taken place.

Business Intelligence: The applications, infrastructure, tools and best practices that enable access to and analysis of information to improve and optimize decisions and performance.

CRUD: This acronym stands for create, read, update and delete. It describes access rights for data.

Change Control: A formal process used to ensure that a process, product, service or technology component is modified only in accordance with agreed-upon rules. Data governance programs often strive to extend the scope of change control to include additions, modifications or deletions to data models and values for reference and master data.

Compliance: A discipline set of practices and organizational group that deals with adhering to laws, regulations, standards and contractual arrangements. It's the adherence to requirements. Data governance programs often support many types of compliance requirements.

Control: A means of managing a risk or ensuring that an objective is achieved. Controls can be preventative, detective or corrective and can be fully automated, procedural or technology-assisted, human-initiated activities. They can include actions, devices, procedures, techniques or other measures.

Data Architecture: A discipline, process and program focusing on integrating sets of information.

Data Custodians: Persons responsible for the safe custody, transport and storage of the data and implementation of business rules. They are responsible for the technical environment and database structure.

Data Dictionary: A centralized repository of information about data such as meaning, relationships to other data, origin, usage and format.

Data Governance: The exercise of decision-making and authority for data-related matters; the strategy of managing and controlling data; the organizational bodies, rules, decision rights and accountabilities of people and information systems as they perform information-related processes. Data governance ensures that data can be trusted.

Data Governance Framework: A logical structure for organizing how we think about and communicate data governance concepts.

Data Governance Methodology: A logical structure providing step-by-step instructions for performing data governance processes.

Data Governance Office: A centralized organizational entity responsible for facilitating, coordinating and supporting data governance and stewardship efforts for an organization.

Data Integration: The practices, architectural techniques and tools for achieving the consistent access and delivery of data across the spectrum of data subject areas and data structure types in the enterprise to meet the data consumption requirements of all applications and business processes.

Data Privacy: The assurance that personal and private information is not inappropriately disclosed. Ensuring data privacy requires access management, security and other data protection efforts.

Data Stakeholders: People who could affect or be affected by organizational data. Data stakeholders include groups who create data, those who use data and those who set rules and requirements for data.

Data Steward: A person with data-related responsibilities as set by a data governance or data stewardship program. Often, data stewards fall into multiple types.

Data Warehouse: A storage architecture designed to hold data extracted from transaction systems, operational data stores and external sources. The warehouse then combines that data in an aggregate, summary form suitable for enterprisewide data analysis and reporting for predefined business needs.

Decision Rights: The system of determining who makes a decision, and when and using what process. Formalizing decision rights is a key function of data governance.

Enterprise Architecture: Enterprise architecture is a comprehensive framework used to manage and align an organization's business processes, IT software and hardware, local and wide area networks, people, operations and projects with the organization's overall strategy.

IT Infrastructure Library: A series of publications providing best practice guidance for IT service management.

IT Service Management: The implementation and management of quality IT services that meet the needs of the business. IT service management is performed by IT service providers through an appropriate mix of people, process and IT.

Issue Framing: A process for scoping and defining a problem before solving it.

Issue Resolution: A structured process for reaching a solution to a problem while considering the needs of all stakeholders. Most data governance programs acknowledge that successful resolution of data-related issues requires politically-neutral facilitation of the decision-making process, with participation by data stakeholders.

Master Data: Master data describes core entities of an enterprise that are used by multiple business process and IT systems; any information that is considered to play a key role in the agency's core operations.

Metadata: Structured information that describes, explains, locates or otherwise makes it easier to retrieve, use or manage an information resource. It is data that provides information about data.

Personally Identifiable Information: Information that can be used to distinguish or trace an individual's identity either directly or indirectly through linkages with other information.

Risk Management: The identification, assessment and prioritization of risks, followed by coordinated and economical application of resources to minimize, monitor and control the probability and impact of unfortunate events.

Sensitive Data: Data that is private, personal or proprietary and must be protected from unauthorized access.

Workflow: The movement of data, documents or tasks through a work process, generally used in the context of technologies that automate workflows. Data governance programs often strive to address workflow by embedding governance controls (e.g., approvals, decision steps) or by providing loop-outs to governance processes (e.g., issue resolution, change control).