



STATE OF INDIANA DATA STRATEGY

Moving Forward Together

Josh Martin, Indiana Chief Data Officer

Version 1.0
February 2024



Executive Summary

Data is a critical asset, and the State of Indiana is dedicated to strategically leveraging the massive volume of key data collected by its agencies throughout the enterprise. Through consistent, innovative usage, the state can deliver the best service to Hoosiers. This is being achieved through several key initiatives in 2024, including:

- * DCAM Capabilities Assessments
- * Artificial Intelligence Policy
- * Enterprise Data Catalog implementation
- * Continued Agency Privacy Officer training
- * Data Proficiency Program expansion

Overview

The Indiana Chief Data Officer is statutorily charged with setting standards for data management and governance in order to achieve the ultimate goal of enhancing the value, quality, security, and understanding of data through coordinated efforts of statewide stakeholders. This enables continued evidence-based decision-making and data-informed policy-making that best serves the citizens of Indiana. For example, data is critical for establishing the foundation for evidence and evaluation necessary to ensure programs are effective and achieving their intended outcomes.

A strong and consistent approach to the State's data challenges will generate benefits to individual agencies and the enterprise overall, including:

- Increased efficiency
- Compliance
- Cost savings
- Improved transparency and accountability
- Improved service delivery

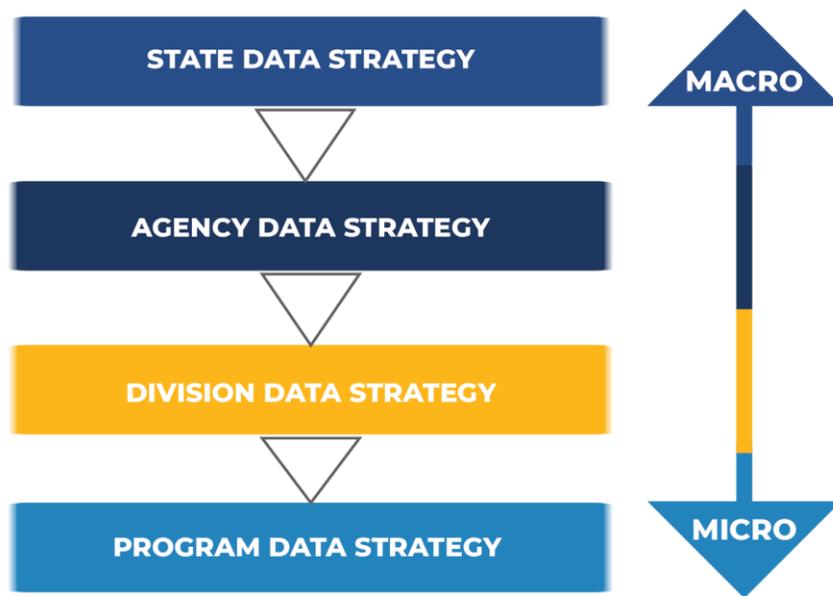


Figure 1: State of Indiana Data Strategy Continuum

The following Enterprise Data Strategy outlines high-level initiatives that will benefit the State of Indiana as a whole, but it represents only one piece of the strategic planning necessary to support a strong data ecosystem. Each agency should have their own data strategy, cascading down into division- and program-level strategies tailored to the types of data and usage specific to each of those operational levels (Figure 1). This approach will ensure that data is treated as an asset from the smallest to largest use cases and governed at each level by subject matter experts.

MPH’s current enterprise initiatives include evaluations of agencies’ current data management/ governance capabilities to create maturity roadmaps, the introduction of an enterprise data catalog, and the establishment of agency privacy officers. These efforts require agencies to look internally at the way they view and use their own data as a way to improve overall data quality, access, and usage. Quality data is the foundation to innovation, ease of use and more robust data initiatives. Enterprise-wide acceptance empowers the State to realize the value of state data as a strategic asset.

Here is a high-level timeline of what MPH expects to achieve through the end of 2024:

Initiative	Measures of Success	Deadline
Data Capability Assessments (DCAM)	10 agencies or business line assessments completed	October 31, 2024
Artificial Intelligence (AI) Policy	Delivered to agencies	February 26, 2024
Enterprise Data Catalog implementation	10 agencies onboarded into Informatica	October 1, 2024
Agency Privacy Officers	78 agency APOs identified and trained	June 1, 2024
Data Proficiency Program Employee Participation	3,000 state employees earning at least one badge	December 31, 2024

Figure 2: Enterprise Initiatives Timeline

To learn more about individual elements, read on below:

State Data Governance Framework

Data Management Capability Assessment Model (DCAM)

The Data Management Capability Assessment Model (DCAM), the State of Indiana's data governance framework, establishes agency baselines to promote moving forward together. Initial assessments enable agency management to determine current state overall and for lines of business to strategically plan next steps. The DCAM Assessment also identifies strengths and weaknesses, compares areas within organizations, and defines quantifiable metrics. These metrics help drive to a better, more defined standard.

The Enterprise Data Governance Catalog (EDGC) provided through Informatica will allow every agency the opportunity to scan their data libraries so that their data is available at their fingertips while operating within relevant privacy protections. The ability to understand the entirety of the enterprise's data assets at a much deeper level creates flexibility to work collaboratively between State agencies. Data can more easily be leveraged for projects and programs with greater speed and accuracy.

Approach

Beyond tooling, an effort must be made to educate and enhance the abilities of the people that interact with the State's data. MPH's Data Proficiency Program, available to all State employees, provides an introduction to data's importance and utility in day-to-day decision-making. Agency leadership is furthermore encouraged to send their lead data expert to attend quarterly Data Governance Council meetings in which critical updates are shared. They may also participate in council subcommittees organized to deliberate on high-priority issues to recommend new statewide policies. Critically, the State must invest in its data privacy and ethics infrastructure through training of designated Agency Privacy Officers.

Data is everywhere. The technology is quickly evolving. The nature of this space means that additional tools and strategies are constantly on the horizon, with new opportunities presenting themselves outside the constraints of a long-term strategic plan. Therefore, some initiatives may be introduced beyond this document's proposals.

Scope

The goal of the State Data Strategy is to provide the high-level framework and guidance for realizing the full potential of data across the enterprise. This strategy is intended to be overarching guidance that will help all agencies work toward best practices in data management across the State. The data strategy is not designed for daily operational management of internal agency data or data programs.

Sparking Innovation

The State of Indiana provides numerous programs and services for Hoosiers. Quality work requires quality inputs, including high-quality data. MPH partners with other State agencies and external partners to translate data into information, bringing together disparate sources to derive the greatest value for our State's data assets. Bringing together resources like data, agencies, technologies, and skilled partners to collaborate results in innovative solutions to the complex issues facing Hoosiers today.

The driving force throughout this data strategy is in the collaborative use of tools that will allow State employees to work together more effectively and efficiently. Across all sectors, everyone is talking about artificial intelligence (AI) and machine learning (ML) as the next big breakthroughs. Few are discussing the importance of having the correct foundational elements in place to empower AI/ML to work correctly, securely, and without bias. This strategy helps the state to put build or strengthen those foundational requirements.

This strategy also incorporates a statewide culture of learning and collaboration while encouraging agencies to model transparency. Education will be a key component of both framework and product implementations so that agencies will better understand the benefits of feedback loops, data sharing, and data governance best practices.

Processes

Assessing Data Management Maturity

State agencies vary in their current data governance maturity level. Establishing a minimum data management standard across the enterprise, based on a common best practices framework, is of paramount importance. Data management involves the development, execution, and supervision of plans, policies, programs, and practices. These in turn deliver control and protection while also enhancing the value of data and information assets throughout their lifecycles. No matter where an agency is in their data journey, the MPH Governance Team is administering DCAM assessments to establish current states and advise on road-mapping efforts.

Action	Description	Deliverables
MPH's Governance Team works with agency leadership to distribute the DCAM assessment for targeted use	Staff at all levels of a targeted agency or line of business will be assessed to establish baseline knowledge of the unit	DCAM Assessment and results analysis at the agency or line-of-business level Digital DCAM Assessment
MPH's Governance Team meets with agency leadership to review the assessment results and schedule a consensus-building meeting	A review of responses establishes next steps for implementing more governance practices in daily agency data activities	A sentiment analysis session of assessment responses with all assessors led by an MPH Governance Team member
Agency leadership implements plans to drive to a better, more-defined standard of managing data consistently	Strong data management and governance policies improve data quality and interoperability	Strategic Plan and Roadmaps
Agency leadership participates in the Data Governance Council to create policies around data governance for agencies and the State of Indiana	The Data Governance Council is made up of agency leadership and is a cross-topic, interdisciplinary council inclusive of information technology and business stakeholders that holistically address data governance and standards	Quarterly Data Governance Council Meetings Additional solicitations for feedback on policies Task forces and focus groups on specific policies
Individual agency representatives work within Data Governance Council sub-committees to standardize policies relevant to their areas of focus	Subcommittees address policy areas, personnel development, authoring standards, hot-button topics, and more	Subcommittee meetings Subcommittee recommendations to the Data Governance Council

Figure 3: DCAM Assessment Process Flow

Enterprise Data Governance Catalog (EDGC) Implementation

An important step in bringing the State of Indiana's data assets together is the implementation of an enterprise Cloud Data Governance Catalog (CDGC). A data catalog is an organized inventory of data assets that relies on standardized metadata templates to optimize discovery and utilization.

Informatica is providing the State's solution for this effort, with MPH leading the implementation for partner agencies. This effort will deliver the ability to discover, organize, relate, govern, and distribute information of the data assets within the enterprise. Captured metadata elements include classification, ownership, authoritative data sourcing, data lineage, data profiles, and quality, entitlement, data access, and usage purpose

tracking, data lifecycle tracking, data protection need, privacy requirements, and residency details. While this will take time to enact across the enterprise, it is a key priority that encourages collaboration, enhances security, and enables robust AI/ML opportunities in the future.

Action	Description	Deliverables
Each agency works with MPH to implement their data catalog	MPH meets with individual agencies to introduce them to the Informatica product, share documentation, and address security concerns around sensitive information	Informatica implemented at agency level Dedicated MPH resources for assistance
Each agency scans their agency data into the Flat Organization	The Flat Organization houses all State data assets and makes their metadata viewable to authorized users to facilitate discovery and sharing across agencies	Agency data scanned into the enterprise Cloud Data Governance Catalog Quick Start/ Usage Guide for Informatica tooling Enterprise standards and policies for use of the CDGC
Each agency continues to scan in any new data as it is acquired and use the quality tools to facilitate bettering the data within the tool	As new data is generated or acquired, it is added into the data catalog	Timely updates to CDGC of new data assets Use of data quality tooling within Informatica product CDGC Data Quality Policy
Data catalogs that are already scanned and in the tool are maintained, updated, or retired as necessary	A current picture of available and up-to-date data is maintained	Updated data catalog Policy outlining life cycle management of agency CDGC assets
Lessons learned are shared within the agency and others throughout the State in designated user groups and the Data Governance Council	User groups within the agency can be created to leverage benefits internally, while enterprise groups share best practices, hints, tips, and tricks to achieve maximum value	Creation of an Informatica User Group Updated FAQs and User Guide

Figure 4: Enterprise Data Catalog Process Flow

Establishing an Agency Privacy Officer (APO)

Data Privacy and Ethics is the piece of data management that addresses the appropriate handling of personal data in compliance with data protection laws, regulations, and general privacy best practices, as well as recommendations of appropriate conduct in relation to data. The State’s Chief Privacy Officer and their team are housed within MPH, setting standards that enable a consistent, comprehensive enterprise approach to protecting sensitive data assets. In an effort to expand and implement a strong privacy and ethics environment across the State, each agency should designate an Agency Privacy Officer to serve as the primary administrator of these policies internally. They will also work across agencies to raise concerns, offer feedback, and generate best practices to strengthen the ecosystem overall. In most agencies, serving as the privacy officer is not a full-time expectation, but an additional duty that is complementary to the individual’s current position. Many are the agencies’ general counsels. To assist, MPH is providing regular curriculum to prepare privacy officers for the role.

Action	Description	Deliverables
All agencies appoint an Agency Privacy Officer (APO)	Each agency designates a single point-of-contact and chief decision-maker for privacy issues	Designated Agency Privacy Officer
The APO is responsible for training their agency members on all state policies regarding Privacy and Ethics	The State’s Chief Privacy Officer (CPO) works alongside Data Governance Council members to provide data standards and processes applicable to all agencies with the assistance of APOs	State Privacy and Ethics Policies Agency training curricula
Annually, the APO evaluates if a Privacy Impact Assessment (PIA) is required	The APO partners with program and system owners to identify new items that would need to be reviewed within their agency, then work with the CPO to determine if the PIA is necessary	Evaluation guidelines Privacy Impact Assessment
The APO educates themselves and follows any state-wide or federal guidelines that could potentially affect their agency	The APO is a resource for their agency, engaging the CPO whenever additional assistance is required	Circulation of applicable state and federal guidelines

Figure 5: Agency Privacy Officer Process Flow

Creating a Mature Data Infrastructure

A cooperative effort is needed to advance the people, processes, and technology within the State to make this strategy work. The State of Indiana Data Strategy continues to evolve along with technology and advancements, but the actions listed above require support and engagement from agency leadership, data experts, and employees at all levels within the State of Indiana in order to be fully realized. The benefit is an enhanced understanding of the State's available data, streamlined processes to ingest and manage new data, and an increased ability to leverage it as a strategic asset in the years to come. A mature data infrastructure requires teamwork. Through effective communication, Data Governance Council and sub-committee participation, and implementing standards across the Enterprise, the State of Indiana will move forward together.