

Introduction to Data Culture

Data culture refers to the collective mindset, practices, and values surrounding the use of data within an institution. It encompasses how faculty, staff, and administrators engage with data to inform decision-making, improve student outcomes, and enhance institutional effectiveness. A strong data culture fosters transparency, collaboration, and trust, ensuring that data is accessible, understandable, and used ethically. At its core, data culture is a shared practice of co-creating meaning from data; starting with its generation, continuing through its application, and culminating in its impact.

The purpose of data culture is not just about technology or analytics, but rather about creating an environment where data is integrated into everyday processes to support student success and institutional goals.

This includes:

- Encouraging data literacy among faculty and staff so they can interpret and apply data effectively.
- Using data to identify trends in student performance, retention, and engagement.
- Establishing clear data governance policies to ensure responsible data use.
- Breaking down silos between departments to promote data-informed decision-making.

A well-developed data culture helps institutions move beyond intuition-based decisions to evidence-driven strategies, ultimately leading to more equitable and effective educational practices.

Data Catalog

The first step in strengthening your institution's data culture is gaining a clear understanding of the scope of your data; how it is created, managed, and utilized. A simple yet effective starting point is to determine whether your institution maintains a data catalog and to identify the unit responsible for its oversight. If no such catalog exists, or if user-facing resources about data assets and availability are limited, initiating outreach and cross-departmental communication can lay the groundwork for developing a comprehensive data catalog.

A data catalog utilizes the details available at an institution to provide a user facing summary of available data in an institution along with details relevant to its use. Data catalogs serve a basic function of informing users while providing a shared understanding of the institutional data availability and general context.

Data Catalog

Data catalogs often begin as simple lists of data sources and 1-2 sentence description about the focusing on the who, what, where, when, and how questions that are pertinent to stakeholders.

- Who can I reach out to about this data? Systems specialist, Unit manager, Data developer, IT professional, External resource
- What is the content of this data? Student, Academic Performance, Faculty Information, Accreditation
- Where can this data be accessed or found? Warehouse, Local drive, Source system
- When does the institution receive this data? Daily, weekly, quarterly, annually
- How does the institution receive this data? API, Manual entry, Download, Reports

Answering basic questions about data helps build awareness about institutional assets, fostering collaboration, and improving transparency. In addition, basic frameworks like data catalogs can support the development of operational documentation, internal metadata development, and programmatic direction.

Glossary

Data Inventory - a structured record of an organization's data assets, detailing what data is collected, where it is stored, and how it flows within the business. It plays a crucial role in compliance, informed decision-making, and operational efficiency by providing visibility into data usage and governance¹. Unlike a data catalog, which organizes and contextualizes data for easy access, a data inventory focuses on technical metadata and system integration requirements.¹

Data Catalog - an organized inventory of an organization's data assets, designed to help users discover, understand, and manage data efficiently. It leverages metadata to provide context, classification, and accessibility, enabling better data governance and analytics.²³

Metadata - structured information that describes, organizes, and provides context for an organization's data assets. It helps users understand the characteristics, relationships, and accessibility of data, enabling efficient data management and governance.

Data Steward - a professional responsible for managing and ensuring the quality, security, and governance of an organization's data assets. They oversee data integrity, enforce data policies, and collaborate with various teams to maintain compliance and usability.⁶⁷ documentation, internal metadata development, and programmatic direction.

References

- 1.Data Science Council of America. (2024, November 29). *What is data inventory and how to create it effectively?* Retrieved from dasca.org
- 2.AWS. (n.d.). *What is a data catalog?* Retrieved from aws.amazon.com
- 3.IBM. (n.d.). *What is a data catalog?* Retrieved from ibm.com
- 4.Herzberg, B. (2022, January 19). *Metadata repositories: Data dictionary vs. data inventory vs. data catalog*. Satori Cyber. Retrieved from Satori Cyber
- 5.Secoda. (2024, September 23). *Understanding the differences between a data inventory and a data catalog*.
.
- 6.Wikipedia. (n.d.). *Data steward*. Retrieved from Wikipedia
- 7.Knight, M. (2024, November 5). *What is data stewardship?* DATAVERSITY. Retrieved from DATAVERSITY
- 8.EDUCAUSE. *Understanding and Developing a Data-Informed Culture*. EDUCAUSE, May 2022,<https://library.educause.edu/-/media/files/library/2022/5/understandinganddevelopingdatainformedculture.pdf>

Authors

Christen White, MS, PMP, Assistant Director, Data Services, UW School of Medicine