Three steps for greater sharing of trusted data

How CDAOs can make data easier to find, understand and access

Table of contents

How CDAOs can make data easier to find, understand and access

"It's estimated that nearly every aspect of work will be optimized by data by 2025—and organizations able to make the most progress fastest stand to capture the highest value."

McKinsey:
Can data improve your organization's operations?

Data can drive innovation and growth, enhance customer and employee experiences, increase efficiency and productivity and improve risk and cost management. Yet the value of data can only be realized if it's easy to find, understand and access by the people who need it, when they need it. Chief Data and Analytics Officers (CDAOs) who enable sharing of trusted data and empower employees across functional areas of the business to consistently use data for decision making, drive better performance and achieve strategic goals more effectively.

Despite the clear benefits of data sharing, organizations face several challenges that can hinder effective implementation. As the volume of data and number of sources accelerate, finding data and determining if it is fit for use becomes more challenging. On average, companies use over 200 applications[1] and the amount of data created in 2025 is projected to be 181 Zetabytes[2]. In addition, organizations must comply with a wide range of privacy and AI regulations that increase the complexity of responsible and compliant data sharing and use.

In the remainder of this ebook, we will examine three key steps CDAOs can take to overcome these challenges to sharing trusted data. By implementing these steps, you can empower broad and consistent use of data across the enterprise and help your business teams deliver more value with data. Let's see how you can share more trusted data, with more people, for more business needs.

1. Okta: Business at Work Study
2. Statistica: Worldwide Data Growth

1.

Simplify data discovery and stewardship

Data discovery helps locate data in today's complex landscapes, while data stewardship ensures that this data is properly classified, cataloged, assessed and governed according to organizational policies and rules. Together, they create the necessary foundation for building an effective and efficient data sharing program.

> ## Nearly one fifth of a data scientist's time is spent discovering suitable data sets.
>
> ACM Computing Survey: Dataset Discovery and Exploration

Discover and classify data

When dealing with tens or hundreds of millions of columns in thousands of data sources – across multi-cloud, hybrid and on-prem environments – it is nearly impossible to manually locate, inspect and evaluate the data. Automation enables you to scale data discovery and classification of data in columns based on categories such as name, credit card number and address, as well as identify domains of data like customer, employee and partner in database records or unstructured data. Documenting lineage and relationships between data can also be automated to further speed the discovery and stewardship process.

Catalog and assess data

A data catalog provides a centralized repository for people to find data assets and determine if they are fit for use. This centralized repository not only facilitates efficient data discovery but also plays a crucial role in maintaining data quality. Automation can support data profiling and data quality assessment at scale based on organizational standards. By integrating data quality into a centralized catalog and displaying data health metrics alongside data assets, organizations can significantly elevate the overall reliability and trustworthiness of their data assets.

Curate and enrich metadata

People will only use data if they understand and trust it. Similar to using automation to find and classify data, you can automate the mapping of data policies, business glossary terms and other useful context to data assets. Furthermore, AI can be used to simplify asset enrichment and reduce the effort needed for time-consuming tasks such as developing asset descriptions. Based on this additional context and other information like certifications and user ratings, data catalogs empower guided data exploration, making it easier for people to find, understand and use data for decision making.

2.

Ensure data privacy and protection

Data sharing programs must balance making data easy to find, understand and access with responsible and compliant data practices. Data privacy and protection reduce the risk of unauthorized use or disclosure of personal and sensitive information.

> ## Nearly 80% of the world's population is covered by a national data privacy law.
>
> IAPP: Global Privacy Laws

Identify sensitive data

Part of the data discovery and classification process should include identifying sensitive data. Automation can help scale the identification of sensitive information such as personal, financial and other regulated information. It can also significantly enhance the process of identifying individuals and entities associated with sensitive information such as customers, employees, and partners, as well as the location of those individuals and entities. Discovering, classifying and labeling sensitive information is the first step toward ensuring robust data protection and compliance with regulations.

Determine regulatory requirements

The regulatory landscape for data privacy and protection is complex and evolving rapidly. National data privacy laws are now in place in 137 countries[3] and governments worldwide are creating and adjusting policies and regulations to achieve the best balance between AI innovation and responsible practices. With broad visibility and a deep understanding of the various types of sensitive information, the associated individuals and entities and their residency, you can effectively map complex regulatory requirements to the data you need to protect.

Map policies and controls

Based on the regulatory requirements, you can map policies and controls to your data assets. Putting policies and controls in place for data collection, consent for use and processing ensures that only authorized people and systems have access to sensitive data and that the processing of data is compliant with relevant regulations such as the General Data Protection Regulation (GDPR). Visibility into policies and controls applied to sensitive data enables auditable data practices that ensure compliant sharing and use.

Identifying sensitive data, determining regulatory requirements, and mapping policies and controls are critical for compliant data sharing and use.

3. IAPP: Global Privacy Laws

3.

Democratize data access

Maximizing the value of data requires as many people as possible to use data in their day-to-day decision making. Data democratization empowers everyone regardless of their technical expertise to feel comfortable and confident accessing and using data.

> ## Less than 50% of organizations say they are data driven.
>
> Wavestone: Data and AI Leadership Executive Survey

Enhance data literacy

Empowering people to understand data terminology, policies, quality and lineage is the foundation of a data-driven workforce. A business glossary is a collection of terms and their definitions that help ensure a common data language and shared meaning. Policies provide guidelines to help people understand how and where data can be used, and quality metrics enable people to determine if the data is reliable. Data lineage helps people understand how data is transformed and used throughout the organization and the downstream impact of changes in data sources.

Deploy data products

Data products are an approach companies are turning to for accelerating data-driven organizational decision-making.

They significantly simplify the process of using data for analysis by people with varying levels of technical expertise by packaging all the functionality needed for analysis such as data schemas, pre-trained machine learning models, user interfaces, governance policies, data contracts and access mechanisms. Some examples include a 360-degree view of the customer, financial forecasting models and product recommendation engines. This approach also makes it faster and easier for people to use data for highly custom use cases by combining existing data products to meet their individual needs without needing in-depth knowledge of data engineering and machine learning practices.

Enable a self-service data marketplace

A data marketplace combines all the work from the previous steps to facilitate internal and external data sharing. Data providers can list their datasets and data sharing agreements/contracts (obligations for both providers and consumers), and data consumers can easily browse, find, assess and request access to desired assets for their use cases. By providing a centralized location with intuitive experiences for diverse data providers and consumers, data marketplaces create a network effect that expands the value potential of data.

Developing data literacy, building data products, and connecting data providers and consumers through a data marketplace empowers a data-driven culture.

CDAO checklist for greater sharing of trusted data

The three key steps explained in this ebook can help CDAOs make trusted data easier to find, understand and access. Using them as a checklist helps you empower everyone in your organization to consistently use data to make decisions. And these best practices help you ensure your data sharing efforts are compliant with relevant internal policies and external regulatory requirements.

1. Simplify data discovery and stewardship

Discovering and classifying data; mapping policies, rules, and terms; and cataloging and assessing data are the foundation for greater data sharing and use.

- ☐ Discover and classify data
- ☐ Catalog and assess data
- ☐ Curate and enrich metadata

2. Ensure data privacy and protection

Identifying sensitive data, determining regulatory requirements, and mapping policies and controls are critical for compliant data sharing and use.

- ☐ Identify sensitive data
- ☐ Determine regulatory requirements
- ☐ Map policies and controls

3. Democratize data access

Developing data literacy, building data products, and connecting data providers and consumers through a data marketplace empowers a data-driven culture.

- ☐ Enhance data literacy
- ☐ Deploy data products
- ☐ Enable a self-service data marketplace

Data is more important now than ever to business success.

It drives innovation and growth, enhances customer and employee experiences, increases efficiency and productivity, and improves risk and cost management. Chief Data and Analytics Officers who enable sharing of trusted data – empowering employees across functional areas of the business to consistently use data for decision making – drive better performance and achieve strategic goals more effectively.

In this ebook we explained three key steps to enable greater sharing of trusted data: Data discovery and stewardship to find and certify data; data privacy and protection to reduce the risk of unauthorized use or disclosure of personal and sensitive information; and data democratization to empower everyone regardless of their technical expertise to feel comfortable and confident accessing and using data.

Are you ready to share more trusted data, with more people, for more business needs?
Check out our guide on four ways you can empower business growth.