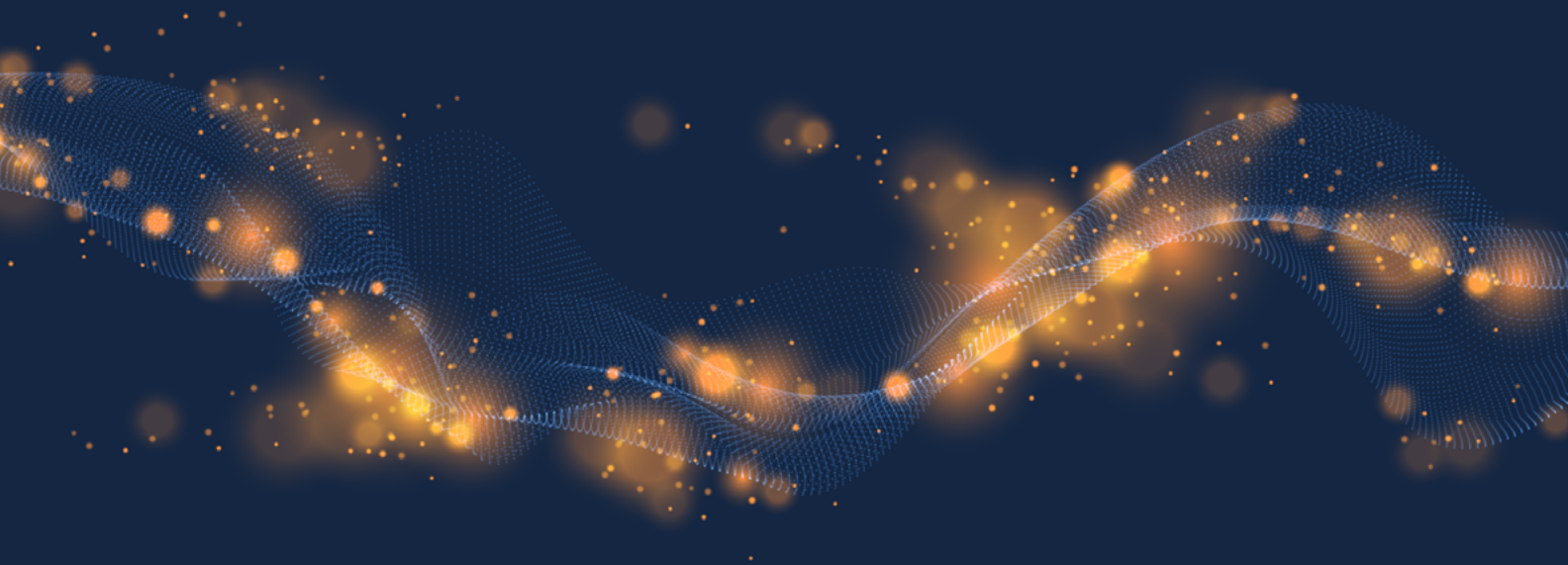


Internal Data Marketplaces

Fern Halper, Ph.D.



Sponsored by

denodo 

INTRODUCTION: WHAT IS AN INTERNAL DATA MARKETPLACE AND WHY IS IT IMPORTANT?

Organizations understand the importance of leveraging data and analytics to remain competitive. As many companies utilize more data for analytics and insights, they want to enrich their data sets for analysis and make data more easily accessible and understandable. Traditionally, organizations have bought third-party data such as demographic or firmographic data to help enrich their own data sets. Over the past three to five years, there has been a move to cloud data marketplaces that facilitate the buying and selling of data. These data marketplaces offer data products such as data sets, industry-specific data products, and analytics outputs.

Concurrent with enriching their data, organizations want to make their internal data more discoverable and understandable. Traditionally, organizations used data dictionaries or glossaries. Over the past five-plus years, many have started to embrace the idea of a data catalog (a central repository that contains metadata for describing data sets, how they are defined, and where to find them) to help them discover, better understand, and better trust their data. Recently, the idea of an internal data marketplace has emerged.

An *internal data marketplace* is a platform where data and analytics within an organization is organized, managed, and made accessible to various business units and teams. It collects metadata and provides a centralized location where users can come and find data. It promotes collaboration and knowledge sharing among teams. This platform provides a marketplace experience for users to come and explore, find, understand, and access data and data products (i.e., derived assets from data) within the confines of the organization.

Some organizations refer to this as a storefront, others as the next-generation catalog or über data catalog.

There are several benefits to using internal data marketplaces for advancing analytics. These include:

- **Centralized data access.** The internal data marketplace provides a single point of access for all data assets so teams don't spend needless time searching for data or re-creating data sets that already exist elsewhere in the organization. With a centralized system, there is also better control and less room for error.
- **Collaborative environment.** Teams can share data sets, analytics, and insights more easily, helping to build a collaborative data culture. They can also endorse data products, ask questions of the data owners, and request enhancements.
- **Data democratization.** An internal data marketplace can make data and data products more accessible to non-technical users, promoting a data-driven culture where employees such as business users can derive insights from data. Ideally, the data marketplace can create data products from any data source in the language of the business, accessible with any technology (SQL, API, messaging, etc.).

- **Data governance, compliance, and security.** An internal data marketplace can help enforce data governance policies such as ensuring that sensitive data is accessed only by authorized personnel and that regulatory compliance is maintained. Internal marketplaces can also help enforce security policies by securing data and data product access with advanced masking and fine-grained security policies to comply with internal policies and regulations.
- **Faster time-to-insight.** With access to the right data in one place, analysts, data scientists, and business users can generate insights more quickly and build data-driven decisions.
- **Timely data.** The data marketplace can ensure everyone is working with the latest version of a data set, avoiding inconsistencies and errors in analyses.

This TDWI Data Points Report discusses the current state of internal data marketplaces. It examines the challenges organizations are facing, the opportunities for internal marketplaces, and the value they can provide.

DATA POINTS SURVEY METHODOLOGY

TDWI Data Points uses primary research to provide a concise view of a specific opportunity area. In August 2023, TDWI fielded a survey targeted at data and analytics professionals. One hundred and forty-three respondents completed the survey and met quality criteria and were included in this analysis. Respondents came from a range of industries, roles, and company sizes.

Survey responses were analyzed and rated across four different dimensions that form the core of the data points analysis.

- **CURRENT STATE** measures how far along respondents are in deploying an internal data marketplace. This is a measure of market adoption. Current state is categorized as early, early mainstream, or mainstream adoption.
- **CHALLENGES** assesses both implementation and ongoing challenges. Overall challenges are rated as significant, manageable, or low.
- **VALUE** evaluates what percent of respondents measured value. Value is categorized as measured impact, perceived impact, or no impact.
- **OPPORTUNITIES** examines opportunities for value. Opportunity is rated as high, medium, or low.

TOP LINE DATA POINTS

Overall Analyst Assessment: It is relatively early for the internal data marketplace and although challenges remain, there is value to be found. (See Figure 1.)

As illustrated in Figure 2, not all respondents to this survey were using an internal data marketplace. Some were using data catalogs. Others were using cloud provider marketplaces. Some were using a combination of platforms. Results are discussed from an overall perspective as well as comparing those who are currently using an actual internal data marketplace (or any marketplace at all) to those who are not.

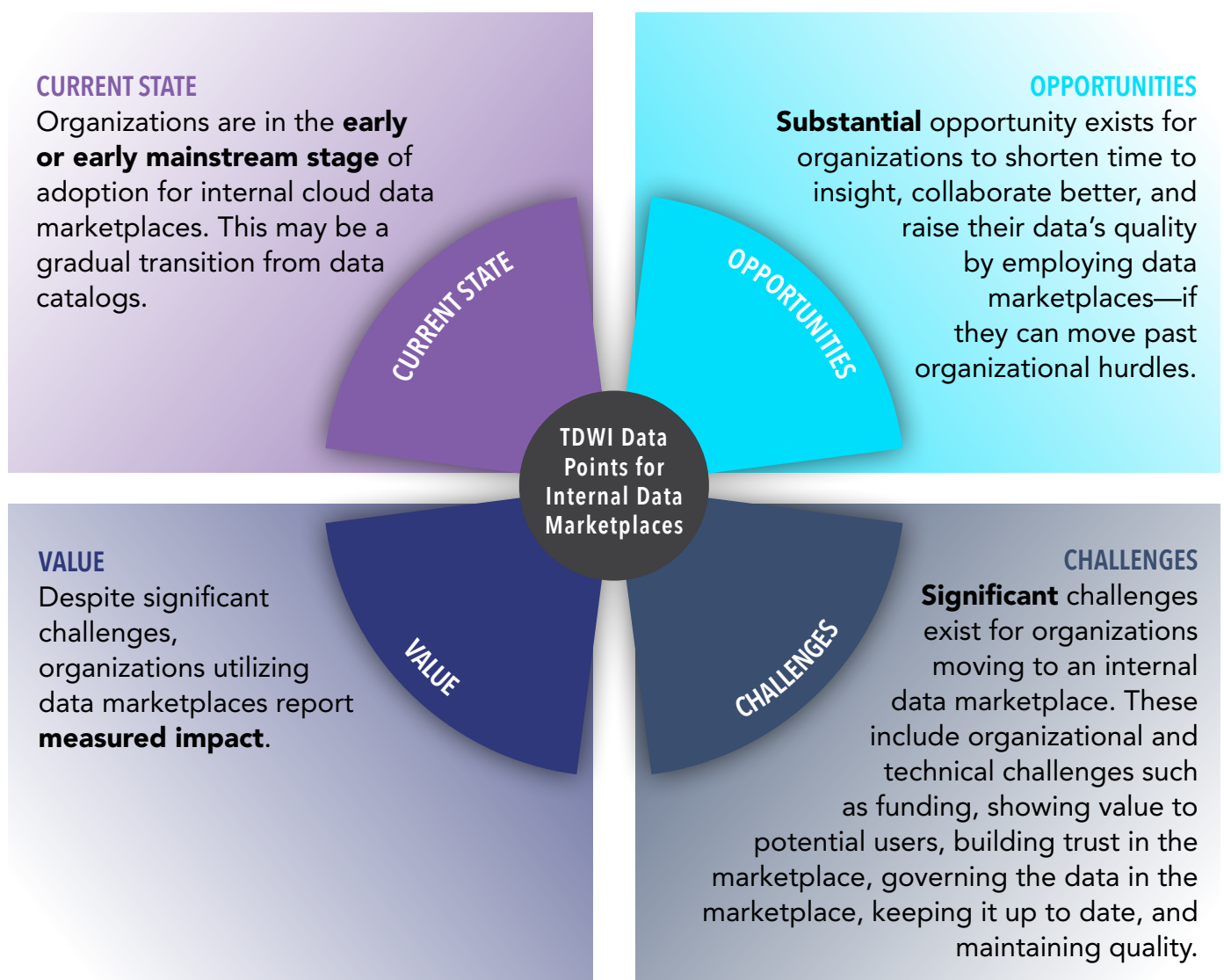


Figure 1. Four key data points for internal data marketplaces.

RESULTS BREAKDOWN

CURRENT STATE

The Data Points survey asked respondents about the current status of their internal data marketplaces. As Figure 2 reveals, 33% of respondents stated that they have deployed an internal data marketplace. Thirty-five percent are deploying data catalogs and 32% are utilizing a third-party cloud provider marketplace. There is overlap among these categories. Close to 40% of those respondents using an internal data marketplace are also making use of a data catalog or a third-party cloud provider's marketplace (not shown). In other words, a subset of those who use internal data marketplaces also make use of other kinds of marketplaces and catalogs. However, 26% of respondents are not making use of any kind of marketplace or catalog. There is no difference in adoption among different company sizes.

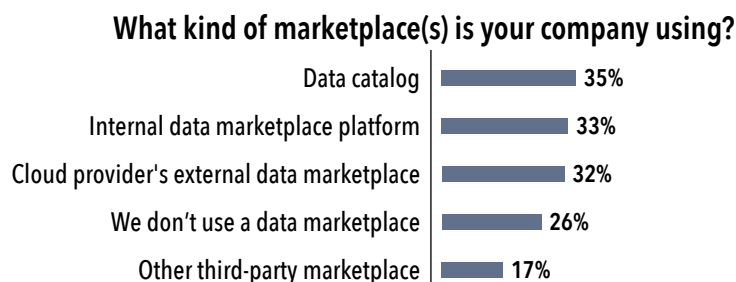


Figure 2. Based on 143 completed responses.

These results suggest that the internal data marketplace is entering the early mainstream phase of adoption.

When asked who was making use of the internal data marketplace or who might make use of it; respondents selected data analysts (66%) followed by business users (65%), and data engineers and data scientists (48% each, not shown). In other words, respondents believe that internal data marketplaces should be used by various personas across the organization to drive value. For those already using an internal data marketplace, business users were the top response (75%, not shown), indicating that this small subset of respondents are on their way to democratizing their data and analytics. A large percent (65%) of this group also reported that data engineers were using the internal data marketplace.

When asked what the internal data marketplace is being used for, responses from those using one included “searching and discovering data assets,” “discovering data sets, understanding their metadata, dependencies,” “creating data sets for analysis,” “making informed decisions,” “building trust in data,” and “sharing insights as data products.”

We asked respondents to the survey about the kinds of data that are or should be available in the data marketplace (Figure 3). Interestingly, although structured data is the top kind of data accessed by those currently using a marketplace (66%), among those not yet using a marketplace, data products (which ranked second overall) would

be the top kind of data asset they would like to access (not shown). Data products are assets derived from data. For example, a data product could be a data set or a dashboard or even an analytics application. These data products could be provided to internal customers as well as partners and even external customers. This is the direction that internal marketplaces are headed. They can provide data as well as analytics products.

If you are using an internal data marketplace, what kind of data is available?

If not, what kind of data would you like to see available?

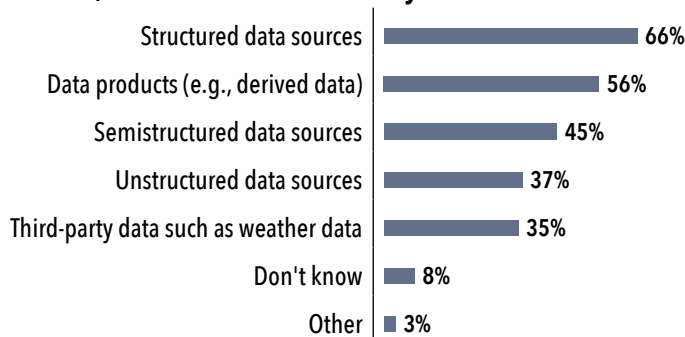


Figure 3. Based on 143 respondents.

CHALLENGES

Respondents reported facing significant challenges implementing internal data marketplaces. For instance, among those currently using a marketplace, not everyone who performs analysis uses it. In this survey, slightly more than half of the respondents stated that those performing analysis access an internal data marketplace (not shown). There are numerous reasons why the rest don't use internal marketplaces, data catalogs, and third-party marketplaces (Figure 4). The top reason is that the organization isn't data-literate enough to use the marketplace (36%). In other words, although others in the organization are analyzing data, they don't have the know-how to make use of a marketplace.

This could be because they haven't been trained on how to use it (32%) or it just isn't in their wheelhouse (perhaps they are just used to using one BI tool). Thirty percent of respondents cited that some in their organization don't find the data in the marketplace to be valuable. This could be because the marketplace doesn't have the data assets they need, the data isn't trustworthy, or too few people in the organization are buying into the marketplace concept.

On the technical front, the top challenge is governance of the marketplace data (45%) followed by keeping the data up to date, data security, and the quality of the data in the marketplace (all 36%, not shown). These challenges are the same for those already using an internal data marketplace. These are issues that we've also seen people talk about with data catalogs.

It is one thing to build a data marketplace and load it with data, but who is going to be responsible for keeping it up to date and ensuring the trustworthiness of the data? This can be a significant challenge if there is no corporate alignment or funding to hire data stewards, data product managers, and the like.

There are also current and perceived organizational challenges with the internal data marketplace. On the organizational front, the top challenge is alignment with corporate and data initiatives (45%), followed by the culture for self-service empowerment (39%) and funding concerns (36%, not shown). Some users are not aware that a marketplace exists. Respondents reported difficulty selling the concept to executives as well as ensuring that people believe the data in the marketplace is trustworthy. As one respondent said, “No governance over self-service leaves it to grow like a weed that then gets killed off.”

Why aren't more people in your organization using the data marketplace?

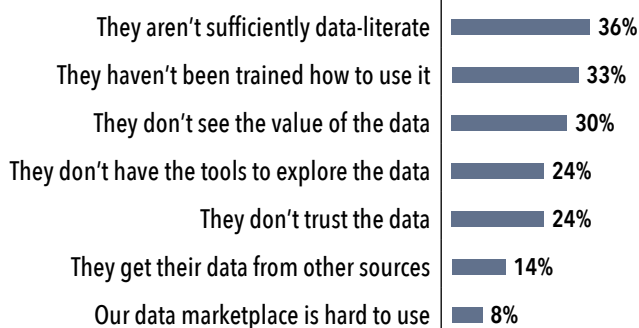


Figure 4. Based on 143 respondents. Up to three responses allowed.

VALUE

We asked survey respondents about what value they have seen from their internal data marketplaces. Although there are real and perceived challenges, respondents did report value. Overall, for those using internal data marketplaces, data catalogs, and external cloud marketplaces, close to 45% claimed to have quantified a top- or bottom-line impact with their internal data marketplace because analytics adoption has increased and analytics-related metrics have improved (not shown). Close to 25% responded that they believe they have gained top- or bottom-line benefits because more users are analyzing their own data, gaining better insights, and doing more with data (not shown). The percentages are even higher when just examining those using an internal data marketplace.

When asked about the benefits of the internal data marketplace, the top benefit was faster time to insight (47%), followed by better collaboration (41%) and better overall insights (38%, Figure 5). For those already using a data marketplace, faster time to insight (51%) and better collaboration (38%) ranked in the top three. However, higher quality data for analytics (43%) was also a top-three benefit (not shown). This suggests that the internal data marketplace can improve the trustworthiness of the data.

What top three benefits have you seen/do you think you will see from an internal data marketplace?

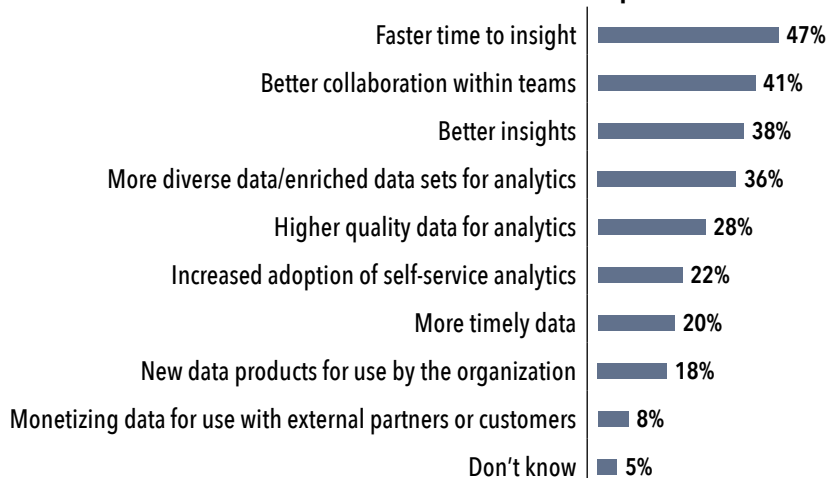


Figure 5. Based on 143 responses. Up to three responses allowed.

OPPORTUNITIES

Although organizations may face challenges with internal data marketplaces, those using them are deriving value. Based on the results of this survey, TDWI sees opportunities for organizations to overcome challenges and gain value by considering the following tools and best practices:

- Communicate the value of the marketplace.** An internal data marketplace isn't a "build it and they will come" product. For organizations to get the most value from it, employees need to know it exists and be told the value it provides. Organizations starting out on the marketplace journey should build a communications strategy for it. Market the launch of the internal data marketplace. Produce videos and other communications material letting potential users know it is coming and how it can help them with their analytics. People don't want to be told to use a marketplace; they need to want to use it.
- Ensure that the right roles are in place.** For the marketplace to succeed, the data will need to be curated and kept up to date. The department that produced the data (the data producers) typically would own that data. For instance, that might be sales, marketing, or an operations team. However, the company may have data stewards in place to manage the quality and accessibility of the data. This role is often in IT. Recently, some organizations are creating a new role called the data product manager with stewardship responsibilities in addition to data "product management" responsibilities.

In some cases, organizations may create new derivative data products from the original data sets and will need to determine whether the ownership becomes shared (between the original producers and those who are deriving new products from the data) or what organizational structure will fit best with the company's policies.

- **Build data literacy.** To make the greatest use of the internal data marketplace, those in the organization need to better understand topics such as metadata, data products, and how to use the actual marketplace. This will require training. TDWI sees a new team emerging, sometimes called the data literacy enablement team. This team is responsible for assessing, executing, and tracking data literacy programs for various personas in the organization. A business user may need a different kind of training for the internal data marketplace than a data engineer, for instance.
- **Build the culture.** Tied to the opportunities already mentioned, a data-driven culture will need to be put in place so the internal marketplace will succeed. This is made much easier if there is someone in the organization in a leadership role who can evangelize the marketplace. Ideally, this person is also part of the C-suite and can influence their colleagues. This will ensure that funding is made available and the company is moving in the right direction.
- **Select the right internal data marketplace platform.** Although overcoming organizational challenges is key to the success of an internal data marketplace, the right technology also needs to be in place. The marketplace must support data from multiple environments because most organizations have a hybrid architecture (some data catalogs may only reflect one source of data). It should enable users to search for data and other products through an intuitive interface. Obviously, the marketplace will need to handle metadata so users can understand the data in the marketplace, what data may be part of an analytics asset, how it was derived, and so on. It will require that approval workflows be put in place so someone seeking necessary data can follow a defined process to get access to that data.

The marketplace must be able to keep data up to date. That means as new data is created in outside systems, the marketplace can reflect that. The data will also need to be kept secure and meet the company's governance policies. The marketplace should enable sharing insights for collaboration. Finally, it will also be important to monitor who is using what data in the marketplace for security as well as performance purposes.

SUMMARY

As organizations move to become more analytically mature, the internal data marketplace can provide value by helping users discover data and other data products. This can lead to faster and better insights, a more collaborative culture aimed at business goals, and better-governed data. As this research points out, the key will be to ensure that both the organizational and technology components are in place to make the internal data marketplace successful.

EXECUTIVE Q&A: INTERNAL DATA MARKETPLACES

To learn more about the benefits and current use of internal data marketplaces, we spoke to Kevin Bohan, director of product marketing at Denodo, to get his perspective.

What is an internal data marketplace?

In short, an internal data marketplace is a centralized location where anyone can explore, discover, and access the data they need to do their jobs more effectively. Consider e-commerce vendors, who have made huge investments to enhance their websites with user-friendly navigation, robust search functionality, personalized product recommendations, intuitive filtering options, and seamless checkout processes, creating a convenient, efficient experience that virtually anyone can use.

Organizations that drive greater data utilization and empower users with self-service data and analytics are leveraging many of the attributes you will find on these online shopping sites to make data available to their users. The similarities in the capabilities and user experience are why organizations have adopted the term “data marketplace.”

What’s driving the interest in such marketplaces?

A recent Gartner survey revealed that 47% of digital workers struggle to find the information they need to perform their jobs effectively. Deploying a data marketplace accessible to anyone within the organization will empower your users to find and access the information they need more easily.

How is an internal data marketplace different from an external data marketplace?

Both external and internal data marketplaces offer features that simplify the process of locating necessary information, but their distinct purposes and emphases set them apart. External data marketplaces emphasize generating revenue through single or subscription-based data transactions; internal data marketplaces focus on promoting data sharing within an enterprise, breaking down barriers to enable the distribution of data products and data access across different organizational units.

How is an internal data marketplace distinct from other approaches an enterprise may already use, such as data catalogs?

A data catalog is a centralized tool that aids organizations in the exploration, comprehension, and administration of their data resources, including data from both internal and external data sources. It functions as a data reference hub within the organization, fostering data governance, enhancing data quality, and facilitating data discovery.

In contrast, an internal data marketplace extends the capabilities of a data catalog by providing a platform for internal departments or teams to share and trade data assets. This marketplace enables users to publish, request, and consume data while providing the catalog with data delivery capabilities to enforce access controls and policies, fostering collaboration and dismantling data silos.

The primary distinction lies in the internal data marketplace's role in actively promoting data sharing and facilitating data delivery, thus encouraging data-driven decision-making and cross-functional data accessibility within the organization.

Data catalogs and data marketplaces can also complement each other rather than remain separate entities. The policies and metadata established in a data catalog should seamlessly integrate into the marketplace, streamlining the initial setup and ongoing management of the system.

Who are the primary users of these marketplaces today? Do you see that expanding to other groups?

The primary users include data analysts, business analysts, data engineers, and various departmental heads or managers. As the data marketplace matures and additional data types and sources are included, additional groups will start to leverage the marketplace, including data scientists and non-technical business users who will benefit from easy access to data for self-service analytics.

Are there particular types of companies that tend to adopt data marketplaces?

Internal data marketplaces are not new. They have been used for years, primarily by larger organizations with dedicated data management teams capable of creating them. Today, with comprehensive solutions that cover all the necessary capabilities, even smaller departmental projects or mid-sized organizations can reap the benefits of data marketplaces.

Who is responsible for building and maintaining an internal data marketplace?

This can vary from organization to organization depending on their size and structure. In smaller departmental projects or mid-sized organizations, a compact cross-functional team may construct the system based on input from end users. However, in larger organization-wide deployments, the following roles and departments may participate:

- The chief data officer (CDO) or data leadership team typically takes a central role in initiating a data marketplace. They establish the strategic vision and goals for data management initiatives.
- The data governance team plays a part in designing data governance frameworks and policies that oversee the data within the marketplace.
- The IT department bears responsibility for the technical implementation of the data marketplace, which includes choosing and deploying the necessary software and infrastructure.
- Data owners or data stewards identify and prepare data assets for market inclusion. These stakeholders hail from a variety of departments.
- Business analysts and data analysts assist in shaping the data marketplace by specifying its requirements, which are driven by business needs.

- End users and data consumers contribute by articulating their data requirements and needs, which guide the content of the marketplace.

Ongoing reviews are essential to keep the data marketplace up-to-date and aligned with organizational objectives. The frequency of these reviews and updates depends on factors such as the pace of data changes, evolving business demands, and the organization's scale.

What sets Denodo's solution to build a data marketplace apart?

Denodo's distinctive logical data management approach facilitates the smooth, real-time integration of distributed data sources without copying or replicating. This is complemented by the Denodo data catalog, which empowers users to explore and utilize data assets from across the enterprise. With AI-driven recommendations, collaboration tools, intuitive search, and support for full data lineage retrieval, users can trust in the data they find, making the Denodo Platform an excellent choice for creating a data marketplace.

ABOUT OUR SPONSOR



Denodo is a leader in data management. The award-winning Denodo Platform is the leading data integration, management, and delivery platform using a logical approach to enable self-service BI, data science, hybrid/multicloud data integration, and enterprise data services. Realizing more than 400% ROI and millions of dollars in benefits, Denodo's customers across large enterprises and mid-market companies in 30+ industries have received payback in less than 6 months.

ABOUT THE AUTHOR



Fern Halper, Ph.D., is vice president and senior director of TDWI Research for advanced analytics. She is well known in the analytics community, having been published hundreds of times on data mining and information technology over the past 20 years. Halper is also coauthor of several Dummies books on cloud computing and big data. She focuses on advanced analytics, including predictive analytics, machine learning, AI, cognitive computing, and big data analytics approaches. She has been a partner at industry analyst firm Hurwitz & Associates and a lead data analyst for Bell Labs. She has taught at both Colgate University and Bentley University. Her Ph.D. is from Texas A&M University.

You can reach her by email (fhalper@tdwi.org), on Twitter (twitter.com/fhalper), and on LinkedIn (linkedin.com/in/fbhalper).

ABOUT TDWI RESEARCH

TDWI Research provides industry-leading research and advice for data and analytics professionals worldwide. TDWI Research focuses on modern data management, analytics, and data science approaches and teams up with industry thought leaders and practitioners to deliver both broad and deep understanding of business and technical challenges surrounding the deployment and use of data and analytics. TDWI Research offers in-depth research reports, commentary, assessments, inquiry services, and topical conferences as well as strategic planning services to user and vendor organizations.

ABOUT THE TDWI DATA POINTS SERIES

This series is designed to educate technical and business professionals about current trends, opportunities, and best practices focused on a specific modern topic in analytics or data management. Research for these reports is conducted via surveys of data professionals, and survey findings are supplemented by the insights and perspectives of TDWI analysts and subject matter experts. Enterprises may use these insights to overcome today's data challenges and set their strategies for the future.

© 2023 by TDWI, a division of 1105 Media, Inc. All rights reserved. Reproductions in whole or part are prohibited except by written permission. Email requests or feedback to info@tdwi.org.

Product and company names mentioned herein may be trademarks and/or registered trademarks of their respective companies. Inclusion of a vendor, product, or service in TDWI research does not constitute an endorsement by TDWI or its management. Sponsorship of a publication should not be construed as an endorsement of the sponsor organization or validation of its claims.

TDWI Research provides research and advice for data professionals worldwide. TDWI Research focuses exclusively on data management and analytics issues and teams up with industry thought leaders and practitioners to deliver both broad and deep understanding of the business and technical challenges surrounding the deployment and use of data management and analytics solutions. TDWI Research offers in-depth research reports, commentary, inquiry services, and topical conferences as well as strategic planning services to user and vendor organizations.



A Division of 1105 Media
6300 Canoga Avenue, Suite 1150
Woodland Hills, CA 91367

E info@tdwi.org

tdwi.org