

WHITE PAPER

The Benefits of Business-Driven Data Discovery

witboost

Chapters

Introduction 1

The Current Enterprise Data Discovery Landscape 1
Enter Data Projects 2
Limitations of Traditional Data Catalogs 2

Understanding Business-Driven Data Discovery 3

Introduction to the Data Marketplace Concept 4
Data Catalog vs. Data Marketplace 4
Benefits of a Data Marketplace 5
Key Features of a Data Marketplace 6
How a Data Marketplace and Catalogs Work Together 6

Key Advantages of a Data Markeyplace 7

Business-Driven Discovery with an Integrated Data Marketpalce Solution 8
Creating a Dynamic Data Ecosystem 8
Benefits of Witboost 10
- Data Project Observability 10
- Organization Around Business Domains 10
- Trust Framework 10
- Impact Quantification 10
- Autonomy in Data Discovery and Utilization 11
- Govern Data by Design 11

What Makes Witboost Unique 11
- Cost and Time Saving

The Future of Business-Driven Data Management 12
Satisfying Every User 12
The Path to True Data Democatization 13



1. Introduction

Although enterprise data is everywhere within and outside the enterprise, its strategic business value is overwhelmingly unknown. This will remain the case without business-driven, agile, and comprehensive data discovery, along with democratized yet compliant/secure access management. Data democratization also cannot exist without everyone having data literacy coupled with the ability to access, process, and exchange it.

Statista has reported that 34% of organizations worldwide have had issues with data discovery, 35% with organizational silos that thwart the sharing of meaningful insights, and 39% with poor data quality. However, the root of these challenges is bigger than data silos and access rights across departments. It's about the ability to contextualize data in ways that can address specific business pains and uncover possibilities. This must happen across the enterprise, sector, and even global marketplace/supply chain.

This white paper explains why a data marketplace is critical to propelling organizations toward a business-centric approach to data discovery.

2. The Current Enterprise Data Discovery Landscape

It's been a common theme for over a decade that traditional approaches to data discovery have struggled to keep pace with the rapid evolution of data volume and complexity. Strategies for constructing pipelines to extract and consolidate data in cloud-based lakes have proven difficult to align with business objectives.

An alternative approach of different teams independently harnessing existing data sources has only been marginally better via the process of technology integration on a case-by-case basis.

The resulting traditional approach of extracting necessary data from sources like lakes and warehouses to craft tailored data sets and feeds via countless pipelines leads to a data architecture that is rife with:

- Incomplete discovery
- Redundancy
- High costs
- Maintenance difficulties and lack of scalability
- Data fragmentation that impairs decision-making and leads to duplicated efforts
- Inefficient manual maintenance, management, and accessibility, which creates distrust between data consumers and producers
- Inadequate regulatory compliance hindered by decentralized data processes, the proliferation of data across various platforms, and isolated data repositories

ENTER DATA CATALOGS

The development of data catalogs has eased the need for time-intensive data procurement and formatting. This approach uses standardized data modules and cohesive technologies to:

- Streamline architecture
- Speed up value realization
- Ensure scalability

LIMITATIONS OF TRADITIONAL DATA CATALOGS

Traditional data catalogs, once the cornerstone for understanding and managing data assets, now present several limitations due to aspects such as:

- **Completeness and quality of metadata**: Which can only be maximized when data, metadata, and code follow the same lifecycle.
- **Increasing data volume and diversity**
- **Access limitations**: Manual metadata entry creates bottlenecks requiring more IT time for updating and maintenance; this restricts access to gatekeepers, thus impeding data democratization.
- **Extended time to insight (TTI)**: Manual update lags result in stale data, which makes timely and accurate insights impossible.
- **Escalating costs**: This is caused by high investment for the initial setup, ongoing maintenance costs, and specialized personnel/tools as a lack of scalability and growing complexity take hold.

While the data catalog can still play an important role, that role must change as data environments evolve. Today's enterprises are beginning to focus on business-driven data discovery methods that prioritize agility and control over complex data infrastructures.

3.

Understanding Business-Driven Data Discovery

In 2024, organizations with maximum enterprise intelligence will be able to capitalize on new opportunities five times faster, according to IDC.

The concept of business-driven data discovery is at the heart of those possibilities via its ability to harness data with a focus on business outcomes rather than mere data management. It's about enabling businesses to be fully autonomous in discovering and consuming data while remaining compliant with governance frameworks.

This form of data discovery prioritizes business relevance, ensuring that the data presented for discovery is high in both quality and business utility. The benefits of this approach are clear:

- Enhanced visibility and guaranteed access to trusted, accurate data
- The ability to combine specific data with other discoverable data
- Improved data literacy
- Increased operational efficiency
- Transparency in data usage
- Much faster and easier discovery phase with less data noise by publishing only business-relevant information
- End-user information navigation driven by business concepts and domains, rather than technology silos (data lakes, SQL query engines, etc.)
- Clear data structure, types, and content of each column (structured and semi-structured)
- Clear data ownership, enabling further questions or access permissions
- Clarity on how to consume the data (information on how to connect to the database), the data projects, or whatever else
- Accurate and perpetually updated metadata relationships

An innovative, business-driven data discovery strategy shift overcomes legacy data discovery limitations. It simplifies enterprise data architecture and empowers organizations to derive actionable insights swiftly by aligning data management practices with immediate and future business needs.

The ability to attain these results relies on an emerging approach to data discovery: the data marketplace.

witboost

4.

Introduction to the Data Marketplace Concept

The distinction between data catalogs and data marketplaces is nuanced, yet significant.

DATA CATALOG VS. DATA MARKETPLACE

Internal Data Marketplace	Data Catalog
Help users to share and exchange data, maximizing the use and value of data	Help users to discover, understand and trust the data, maximizing the relevance of the information
Finalize the exchange of data and value	Has an informative function
Only shows business relevant and consumable Data (what you see is shoppable)	Show the entire data asset inventory of the company
Hide unnecessary complexity and information for Data Consumers, improving high level discovery	Allow drill-down on technical information as Data Lineage, reconstructing the full process behind the data
Activate semantic and knowledge to provide suggestions based on user behaviours and characteristics	Allow to create semantic and knowledge and link it to the data
Data must be governance compliant before being published	Is a tool to run Governance
No Data Stewardship	Data Steward is one of the main actors

A **data catalog** is a curated inventory, focusing on data comprehension and governance. It maximizes the relevancy of and trust in data by facilitating its discovery and comprehension. Its informative function shows the entire data asset inventory of the company. Users can drill down on technical information, such as Data Lineage, to reconstruct the full process behind the data. This allows users to create semantics and knowledge that link to the data.

While data catalogs are invaluable for understanding data in its technical context, they often fall short of making data accessible and usable for non-technical business users. A major reason for this is that they require Data Stewards as the main actors.

This is where the concept of a data **marketplace** steps in, marrying the depth of a data catalog with the breadth of an organizational data-sharing platform. Essential functionalities of a data marketplace include role-specific features for consumers, providers, and administrators, as well as overarching functions such as metadata management and privacy security. An internal data marketplace constitutes a dynamic platform where data is not just listed but actively exchanged, with a focus on business-driven, democratized data consumption and application within the enterprise and for selected third parties.

BENEFITS OF A DATA MARKETPLACE

Business-driven data discovery through an internal data marketplace helps users share and exchange data to maximize the use and value of the data. This brings data out of the shadows of technological silos and into the spotlight of business innovation to finalize data exchange and value.

A key characteristic of a data marketplace is that it only shows business-relevant and consumable data —what you see is consumable. This hides unnecessary complexity and information from the data consumer, which enables a high level of discovery.

The internal data marketplace also activates semantics and knowledge to provide suggestions based on usage behaviors and characteristics. Most importantly, a marketplace ensures data is governance-compliant before being published, with no need for a data steward.

The result is a more empowered organization, where data consumers, stewards, and engineers collaborate seamlessly, driving productivity and fostering a truly data-driven culture.

witboost

KEY FEATURES OF A DATA MARKETPLACE

A data marketplace facilitates business-driven data discovery by bringing:

- Advanced search capabilities
- Contextual guidance
- Transparent processes
- Automated data access

These collective features make sure dathat everyone consuming data can efficiently locate and access it while also trusting and understanding it to enable the best business outcomes.

HOW A DATA MARKETPLACE AND CATALOGS WORK TOGETHER

The goal of the marketplace is not to replace catalogs and other query engines but to bring order and a holistic approach to complex environments. The right data marketplace solution can and should integrate concepts such as the data mesh and data fabric—decentralized and technology-driven approaches, respectively, for data management within an organization.

This is also true with the data space concept, which extends beyond internal boundaries to a cross-organizational data-sharing ecosystem.

Such an integrated data marketplace helps organizations:

- Automate the harvesting process for data asset attributes
- Discover all data across every domain in a distributed environment
- Keep metadata updated in real time for both data at rest and in transit
- Assign data ownership and easy-to-understand health and quality scores to facilitate use by any internal group or users

This brings a tremendous amount of power, control, and agility through automation and integration to existing data architectures, making them scalable, cost-effective, and business-driven. However, these are only the broader benefits of a data marketplace.

witboost

5.

Key Advantages of a Data Marketplace

The ability to harness business insights from data analysis is foundational to business innovation, competitiveness, and agility. It is also key to opening and maximizing gains from new markets, products, and approaches to meet growing consumer demands.

The data marketplace provides a one-stop view of data projects and project flows across diverse data sets, which results in:

- Ability to smoothly gain access, and start consuming the data
- Ability to readily query data and avoid manual processes
- Access to a centralized metadata repository that facilitates sharing and collaboration
- Accelerated insight generation
- Independent and efficient data consumption
- Significantly less time and fewer resources required for data-related initiatives
- Lower maintenance costs through integrated marketplace solutions

A data marketplace provides a flexible, scalable solution that increases an enterprise's competitiveness through business-driven discovery, analysis, and decisions.

As with all technologies and approaches, data marketplace solutions often vary in terms of comprehensiveness, complexity, integration, and their overall ability to quickly translate data into business-driven data discovery. However, it’s now possible for an enterprise to get what it wants and needs—including what it didn’t even know it needed—in a single integrated data marketplace solution.

6.

Business-Driven Discovery with an Integrated Data Marketplace Solution

As organizations pivot from data storage and integration to the creation of value through data, the concept of data projects takes center stage.

Unlike data catalogs, a marketplace using business-driven data discovery principles enables organizations to structure data sets into customized, user-targeted products. These would include complete descriptions, access terms, distribution methods, and delivery guarantees, all of which empower stakeholders to:

- Populate their marketplace with their own data projects as opposed to the limitations of third-party marketplace platforms and data mesh limitations
- Identify data opportunities with data project observability across domains and dependency configurations
- Find business opportunities before modeling

CREATING A DYNAMIC DATA ECOSYSTEM

In a vibrant, bidirectional data ecosystem, domain owners become both producers and consumers by engaging in a marketplace that promotes data sharing and collaboration. This ecosystem nurtures a community that values data as a shared asset.

This packaging process ensures that data producers consider the broader utility of their data, overcoming the introspective tendencies of data mesh frameworks and enhancing the value for stakeholders and users across a distributed enterprise. The true essence of an integrated data marketplace is its ability to transform the raw potential of data into tangible business opportunities.

An integrated data marketplace, such as Witboost, transcends the limitations of third-party platforms and traditional data mesh models by offering a bespoke, organization-specific ecosystem. It empowers enterprises to populate their marketplace with proprietary data projects, fostering a culture of innovation and agility.

Witboost revolutionizes data handling by formalizing the concept of a data project. This method requires:

- **Detailed descriptions and metadata**: Providers must detail their data, ensuring clarity on what is being offered
- **Access and licensing**: Clear terms on who can use the data and under what conditions, including distribution methods and service-level agreements
- **Collaboration**: A structured platform for dialogue between data providers and consumers, allowing for customization and innovation in data projects

This streamlines operational efficiencies by:

- **Automating distribution**: Creating pipelines that handle the ingestion, transformation, creation, and delivery of data projects
- **Facilitating consumption**: Allowing consumers to build pipelines that integrate data projects seamlessly with internal systems
- **Tracking use**: Monitoring data consumption to support future product planning and data architecture needs

The ideal business-driven data discovery marketplace delivers functionality that simplifies the publication of data projects, catering to both internal and external partners. The goal is to streamline the end-to-end relationship between data producers and consumers, where:

- **Data producers** can automate data ingestion, cleaning, updating, product creation, documentation, and distribution.
- **Data consumers** can easily discover highly relevant data projects, examine their attributes and ownership, and set up data feeds for integration into target environments.

Leveraging automation simplifies the complex interplay between data management and business application. The ability to ensure businesses can identify and capitalize on opportunities with speed and precision is just the beginning of an integrated data marketplace.

BENEFITS OF WITBOOST

Witboost stands out for its ability to endow business stakeholders with full autonomy to discover, identify, and sample relevant data before consumption. Several key advantages underpin this autonomy.

Data Project Observability



Witboost provides a holistic and panoramic view of data projects across various domains, enabling stakeholders to visualize dependencies and configurations. Multiple platforms and technologies can offer the same business information, which enhances access and understanding while facilitating strategic planning and execution.

Organization Around Business Domains

By organizing data around business domains, Witboost ensures data is accessible, reliable, and contextualized, allowing enhanced comprehension and strategic application. The ability to package data for business domains ensures data assets are curated with the business context in mind for maximum data relevance and noise reduction during the discovery phase.

Trust Framework

Accessible data samples establish a trust framework, assuring stakeholders of the data's integrity and relevance.

witboost

Impact Quantification

Businesses can measure the impact of data initiatives based on specified business outcomes.

Autonomy in Data Discovery and Utilization

Users have the freedom to swiftly and independently identify and leverage the most relevant data sets with advanced search capabilities and user-friendly data visualization; this ensures the discovery phase is both efficient and user-driven.

Govern Data by Design

Witboost emphasizes governance from inception, ensuring that the data within the marketplace adheres to regulatory compliance and internal standards without sacrificing ease of access. This is representative of a key functionality of Witboost, which is its Access Control for streamlining the data access process.

WHAT MAKES WITBOOST UNIQUE

Witboost's technology-agnostic framework ensures data is organized by business domains, not by layers, platforms, or technologies. This ensures maximum reliability for users, allowing for governance by design through automated data project access.

Companies can build a marketplace that not only fosters trust and reliability but is also compatible with any technology stack. Witboost's intuitive interface empowers business stakeholders, data specialists/analysts, and end users to navigate through only the most relevant data with unparalleled autonomy.

Another unique aspect of Witboost is its comprehensive marketplace customization. For example, a company can add micro frontends to integrate additional information sourced from observability, billing, audits, etc.

Cost and Time Savings

Witboost delivers significant cost and time savings:

- **Substantial maintenance cost reductions** compared to traditional data management and discovery approaches/systems
- **25% drop in time to insight** versus the original time frame for data projects
- **Independence in data consumption** due to automated data catalog integration, allowing users to consume data without technical barriers or prolonged wait times

The ideal data discovery marketplace will focus on fostering trust, reliability, democratization, and compliance, which is what organizations that prioritize data integrity and security demand. That is why business-driven data management is the future of every sector from the utility, automotive, and manufacturing sectors to BFSI, retail, and beyond.

witboost

7.

The Future of Business-Driven Data Management

Data will only continue to grow and become more complex as it is gathered from increasingly diverse endpoints. This will require hybrid, end-to-end data management via an integrated business-driven data discovery marketplace. The ideal solution must be capable of adapting to present challenges and needs while acting as a foundation for future innovation in data management.

The stakes are high, with data-driven enterprises averaging 30% growth per year, according to Accenture. Although data sharing is an important part of this growth, the ability to find and verify trusted data sources is only possible with fewer than 5% of data-sharing programs as of the end of 2022, according to Gartner.

Cost savings, security, and automation are equally important in data management across an enterprise, and no business can afford to compromise.

SATISFYING EVERY USER

Witboost delivers an integrated, business-driven data discovery marketplace solution that meets the needs of stakeholders across the business, finance, and technology C-suite while also meeting the requirements of data/analytics professionals and internal end users alike.

The present and future of data is far more than just a byproduct of business processes. It is the central axis around which these processes revolve. In the unfolding narrative of data-driven enterprises, platforms like Witboost will serve as critical facilitators, propelling businesses toward a horizon where data empowers real-world decision-making.

witboost

THE PATH TO TRUE DATA DEMOCRATIZATION

Today's business-driven data ecosystems must foster collaboration and innovation for organizations to compete and thrive. Witboost is at the forefront of this transformation to streamline data projects and democratize data access, rooted in the power to govern data by design instead of manual processes.

As a result, any organization in any sector can create a marketplace built on compatible integration with any technology, fostering trust and reliability.

To learn more about how Witboost can transform your data
through business-driven data discovery capabilities,
click here
and get a quick video demo here

About Witboost



Witboost is a pioneering platform **streamlining complex data projects across various platforms**, enabling seamless **data production and consumption**. This unified approach empowers you to **fully utilize your data without platform-specific hurdles**, fostering smoother collaboration across teams.

It seamlessly blends business-relevant information, **data governance processes**, and **IT delivery**, ensuring technically sound data projects aligned with strategic objectives. **Witboost facilitates data-driven decision-making** while maintaining data security, ethics, and regulatory compliance.

Moreover, **Witboost maximizes data potential through automation, freeing resources for strategic initiatives.** Apply your data for growth, innovation, and competitive advantage.

Find out more at: www.witboost.com

witboost