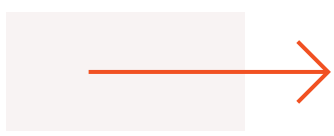


How the Data-Powered Enterprise Unifies and Monetizes Data

Unlock New Business Value with Databricks and Trendence



Maulik Divakar Dixit
Director and Sr. Azure, Databricks and Cloud Architect
Trendence Inc.



David LeGrand
Senior Director, Partner Development and Platform Evangelism,
Trendence Inc.



The Enterprise Quest: Become More Data-Driven

ChatGPT, Claude, Bard. The race to become data-driven has accelerated with the launch of large-language models (LLMs). With each release, these platforms are becoming more powerful. For example, OpenAI's ChatGPT-4 has one trillion parameters, whereas its predecessor, ChatGPT-3, has 175 billion.¹

As a result, executives are discussing how best to deploy discriminative and generative AI to improve workforce productivity, unlock faster innovation, streamline operations, and enhance customer satisfaction. To accomplish this goal, chief data officers (CDOs) and chief information officers (CIOs) strive to democratize data, necessitating a centralized source of trustworthy, quality data so that teams can create trusted analytics, make better decisions, and uncover new opportunities.

In this eBook, we discuss enterprises' current challenges with data democratization and how they can be solved using Databricks' Unity Catalog and Tredence's UnityGO!, an accelerator purpose-built for speeding and scaling migrations to a Unity Catalog enabled Data Intelligence Platform.

Unity Catalog unifies data and enables secure data sharing to downstream systems, while UnityGO! simplifies architecture migration to a Unity Catalog-enabled Data Intelligence Platform. It accomplishes this job by automating the conversion of metadata, such as mountpoints, tables and views; code, in the form of notebooks; and ACLs to Unity Catalog, speeding and simplifying migration by 60% or more compared to manual processes.

Chief Data Officer Priorities for 2023



The race is on to create analytics that enable data monetization through value-producing use cases like predictive analytics that automate processes or customer data platforms allowing data sharing in this cookieless world. CDOs need to improve data quality, access, and governance to accomplish that goal.

CDOs' top priorities include:²

1

Establishing clear and effective data governance (51%)

4

Building and maintaining business intelligence capabilities (36%)

2

Improving data quality (48%)

5

Data monetization capabilities (21%)

3

Building and maintaining advanced analytics capabilities (42%)

6

Data, analytics, and AI ethics (21%)

1. "GPT-4 has more than a trillion parameters - Report," article, The Decoder, March 25, 2023, <https://the-decoder.com/gpt-4-has-a-trillion-parameters/>

2. Sara Brown, "Survey details data officers' priorities, challenges for 2023," article, MIT Management Sloan School, February 21, 2023, <https://mitsloan.mit.edu/ideas-made-to-matter/survey-details-data-officers-priorities-challenges-2023>



60%

Of companies with more than 1K employees have at least partially established data mesh principles.³

3. Martin Whyte, "Data mesh – The next-generation enterprise data platform?," study, 2022, PwC



Why Enterprises Are Governing and Democratizing Data

Democratizing data access improves internal collaboration but comes with a burden. Data must be governed and secured to enable trustworthy sharing and collaboration and increase users' decision-making confidence. Internal teams gain easy access to a centralized source of ready-to-use data for analytics and artificial intelligence (AI) initiatives, commonly called "one source of truth." As an example that is relevant to today's cookieless world, external partners can securely experiment with data they would not commonly have access to. They can build retail media network data cleanrooms that lower marketing costs and provide granular insight into customer preferences and behavior..

The benefits of democratizing data with Databricks and Tredence include:

■ Centralizing data

With Unity Catalog, teams can still maintain specialized data applications for specific functions but also can access and consolidate mission-critical data on a single platform. Databricks Data Intelligence Platform provides one platform for all data, analytics, and AI activities, no matter where they reside.

Use case example

Key transactional purchase data in a propriety data warehouse is combined in the Data Intelligence Platform with supply chain product costing data to provide granular cohort insights for collaborative price planning.

■ Gaining a single source of data truth

Databricks Data Intelligence Platform uses an open-source architecture to integrate structured and unstructured data, enabling teams to use all internal and external data for analytics and AI value creation.

Use case example

External weather data is combined with fuel consumption data to optimize delivery route efficiency.

■ Accelerating processes

Teams provision their data and analytics and access prebuilt machine learning (ML) models on the Databricks Data Intelligence Platform. Tredence's solution-led approach has resulted in the deployment of dozens of industry accelerators built on the Data Intelligence Platform that speed time to value by 50%.

Use case example

Tredence provides over 140 prebuilt use case accelerators across regulated and unregulated verticals.

■ Reducing costs and risks

Using one data platform improves quality, streamlines management processes, and reduces the total cost of ownership. Implementing governance reduces risks, such as duplication errors, corrupt pipelines, or unauthorized access to data.

Use case example

Tredence-built accelerators that enable granular product matching reducing retailers' stock-level inventory costs by 22%.



\$12.9M

Organizations' average financial losses due to poor data quality.⁴



■ Enabling data monetization opportunities

Enterprises want to securely share data to improve first-party data insights segmentation and personalization strategies. They can also sell anonymized data to partners. The data monetization market is expected to grow from \$2.9 billion in 2022 to \$7.3 billion by the end of 2027.⁵



82%

Of data monetization returns are driven by improving operations, while 18% are created by selling data.⁶



The Growing Data Governance Challenge

Data governance challenges are increasing as enterprise data holdings grow exponentially and more functions and users provision and use these vital resources.

Enterprise risk increases when teams duplicate data sets and recycle them across use cases and platforms or when one user corrupts data pipelines feeding multiple use cases. These substandard processes also increase data complexity, cost, and business inefficiencies.

Although Databricks Data Intelligence Platform provides a proven pathway to creating one source of data truth, many enterprises still deploy and pay for multiple platforms, providing teams with a fragmented view of data and insights, resulting in inaccurate decisions based on incomplete data. In addition, failing to apply granular access controls across data platforms increases the risk of its exposure. The average cost of a data breach was \$4.45 million in 2023.⁷

4. Manasi Sakpal, "How to Improve Your Data Quality," article, Gartner, July 14, 2021, <https://www.gartner.com/smarterwithgartner/how-to-improve-your-data-quality>

5. "Data Monetization Market by Component (Tools and Services, Data Type (Customer Data, Financial Data), Business Function, Deployment Type, Organization Size, Industry Vertical," report synopsis, Markets and Markets, May 2023, <https://www.marketsandmarkets.com/Market-Reports/data-monetization-market-127405959.html>

6. "Data is Everybody's Business: The Fundamentals of Data Monetization," article, MIT Management Sloan School, October 16, 2023, <https://mitsloan.mit.edu/press/data-everybodys-business-fundamentals-data-monetization>

7. Douglas Bonderud, "Cost of a data breach 2023: Financial industry impacts," article, Security Intelligence, August 30, 2023, <https://securityintelligence.com/articles/cost-of-a-data-breach-2023-financial-industry/>

The Need for Data Governance

CDOs and CIOs are implementing governance to create secure, scalable solutions that enable better collaboration and data sharing.



400 Average number of data sources per organization.



63% Monthly growth of data volumes.




37% Of data that is currently cloud-enabled.⁸


Generative AI is fueling this fire as enterprises rush to deploy LLMs and domain-specific models, which require extensive amounts of clean data and create governance challenges. New solutions must implement guardrails to ensure that intellectual property and sensitive data, such as personally identifiable information (PII), aren't shared.


As a result, current processes aren't sustainable. **Enterprises need a single solution to integrate all data, analytics, and AI activities to scale processes and reduce costs and risks.**


Introducing the Databricks Unity Catalog


Leading companies have made the smart and sustainable decision to migrate and modernize their data using the Databricks Data Intelligence Platform. Now, Databricks' Unity Catalog creates a universal, governed catalog of data assets on the Databricks Data Intelligence Platform. In simple terms, it delivers:

- 

A single permission model for data and AI
Teams can discover and classify structured and unstructured data, ML models, notebooks, dashboards, and arbitrary files on any cloud.
- 

Data Federation
Users can consolidate, map, and query data from various platforms, including MySQL, PostgreSQL, Amazon Redshift, Snowflake, Azure SQL, Azure Synapse, and Google's BigQuery in one place.
- 

Open Data Sharing
Easy to share data with multiple downstream applications in a secured manner using delta sharing capabilities.
- 

AI-powered monitoring and observability
Unity Catalog has built in system tables to enable auditing and observability into use of the Data Intelligence Platform.
- 

A powerful boost to user productivity
Teams can securely search data, easily understand analytics results, and extract insights from data and AI using natural language queries.

8. "Matillion and IDG Survey: Data Growth is Real, and 3 Other Key Findings," blog, January 26, 2022, <https://www.matillion.com/blog/matillion-and-idg-survey-data-growth-is-real-and-3-other-key-findings>

Teams can reuse the same data, creating consistency across processes and functions. Unity Catalog provides a single permission model for all data and AI, open data sharing, and AI-powered monitoring and observability to streamline critical processes and improve data security.

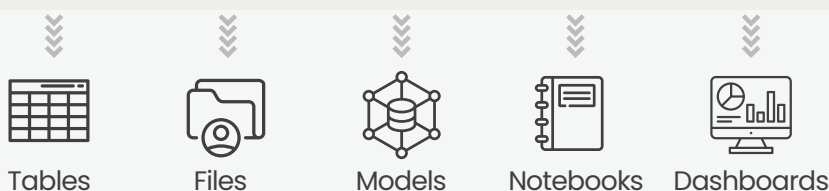
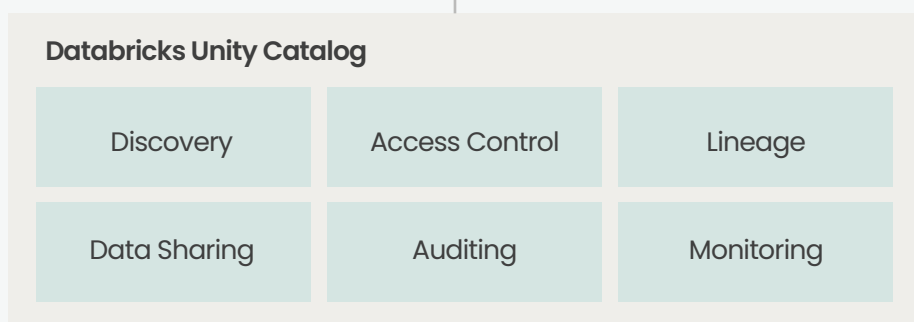
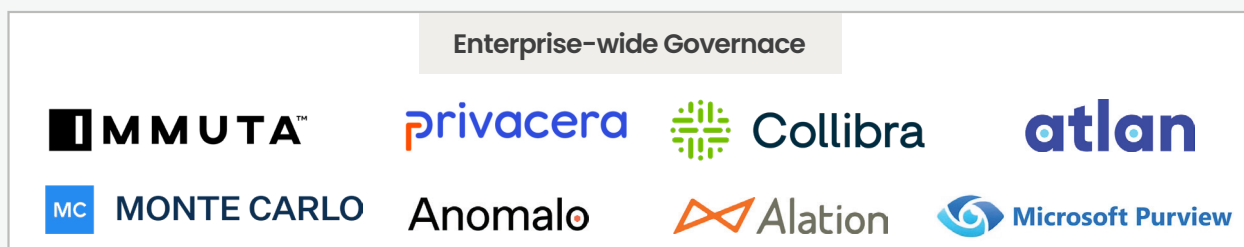
Realizing the Potential of Unity Catalog

Databricks and Tredence have co-developed UnityGO! an AI-enabled tool that enables teams to automate implementation to a Unity Catalog-enabled Data Intelligence Platform. UnityGO! significantly reduces the time and effort (by 60%) to get started with Unity Catalog when compared to manual data engineering migration methods.

By implementing Unity Catalog enabled Data Intelligence Platform with UnityGO!, teams can exchange data with customers, partners, and suppliers while maintaining governance, compliance, and security.

UnityGO!
enables teams to migrate data quickly, reducing time and effort by as much as **60%**

See How Unity Catalog Simplifies Data Governance



Benefits of Deploying Unity Catalog

So what can enterprises expect to achieve when they deploy Unity Catalog? Benefits accrue enterprise-wide across roles and business functions.

CDOs/CIOs



Decrease risk

By centralizing data governance, these leaders end the era of data copies, inconsistent access controls, and outsized security risks due to poorly managed data.

Reduce cost

Data can be shared from any platform and format. As a result, there's no need to rip and replace platforms. Enterprises can maintain specialized platforms to harness their unique capabilities while benefitting from centralized data self-service.

Promote open data sharing

Enterprises will leverage the Unity Catalog to use LLMs, data clean rooms, and other technology as they spur team productivity and innovate with partners.

Security organizations



Improve access controls

Security teams simplify processes using a unified interface to define, apply, and audit access policies for all data and AI assets on any cloud or data platform. They quickly discover and classify structured and unstructured data, machine learning (ML) models, notebooks, dashboards, and arbitrary files on any cloud.

IT/data organizations



Improve model quality

IT and data teams use embedded AI capabilities to automate monitoring of ML models, diagnose errors, and ensure data and model quality.

Business teams



Gain holistic visibility

Business users gain comprehensive observability into Data Intelligence Platform data and AI. They use insights to improve data quality and optimize processes, such as billing, auditing, lineage, and more.

Streamline data access

Business teams use automated migration processes to self-provision the data they need for experiments, speeding the time to results.

Accelerate AI progress

By simplifying the development, use, and maintenance of generative AI, discriminative AI, and ML models, and real-time analytics, enterprises make data-driven decisions and automate processes. By so doing, they run their business with incredible precision.

Use UnityGO! to Unlock the Power of the Data Intelligence Platform

UnityGO! is a purpose-built, Databricks-certified accelerator for Unity Catalog deployments, it delivers the following value to the migration pathway:

1

Lessens risk of unoptimized model/notebook deployment by providing a test workspace environment used to transform to a Unity Catalog enabled Data Intelligence Platform

2

Optimizes migration resources and project planning by assessing workspace migration complexity early in the project allowing for more granular planning and resourcing

3

Improves Databricks native solution (UCX) by augmenting automation to mount point conversion, notebook conversion, moving objects to catalogs and workflow automation

4

Lessens technical resources needed by automating conversion of analytical and AI/ML applications to run on unity catalog enabled Data Intelligence Platform

Teams that use UnityGO!



Automate processes

With highly automated processes, teams save time, reduce business disruption, and accelerate impacts.



Leverage code designed for Unity Catalog

UnityGO! identifies patterns and converts code, automating the creation of workflows.



Reduce risks

Unity Catalog retrofit scripts are tested in a sandbox before being automated. As a result, there's no risk of migration failures.



Can customize patterns

Teams can configure UnityGO! for company-specific patterns and applications, retrofitting existing patterns to run on Unity Catalog.

UnityGO! uses a four-step process to migrate data from any platform or source

1. Setting up the catalog

Teams create a catalog and schema in the Unity Catalog metastore and create objects that point to where data is stored. Next, they create a catalog for development, quality assurance, and production teams in the catalog name.

2. Enabling UnityGO!

Tredence's migration accelerator requires a solution workspace that stores the migration kit that collects the metadata from application workspaces to enable migration. The infrastructure needs to be provisioned for a Web application (a cloud-based service to run web apps and Postgres database) that powers a user interface (UI) that simplifies the migration to Unity Catalog.

3. Collecting and configuring metadata

Teams select the workspaces that need to be migrated to Unity Catalog. UnityGO! automates collection of metadata to help assess complexity of the workspaces and enables configure mountpoints, schemas, managed tables, notebook folders, and user groups in UnityGO! UI Interface.

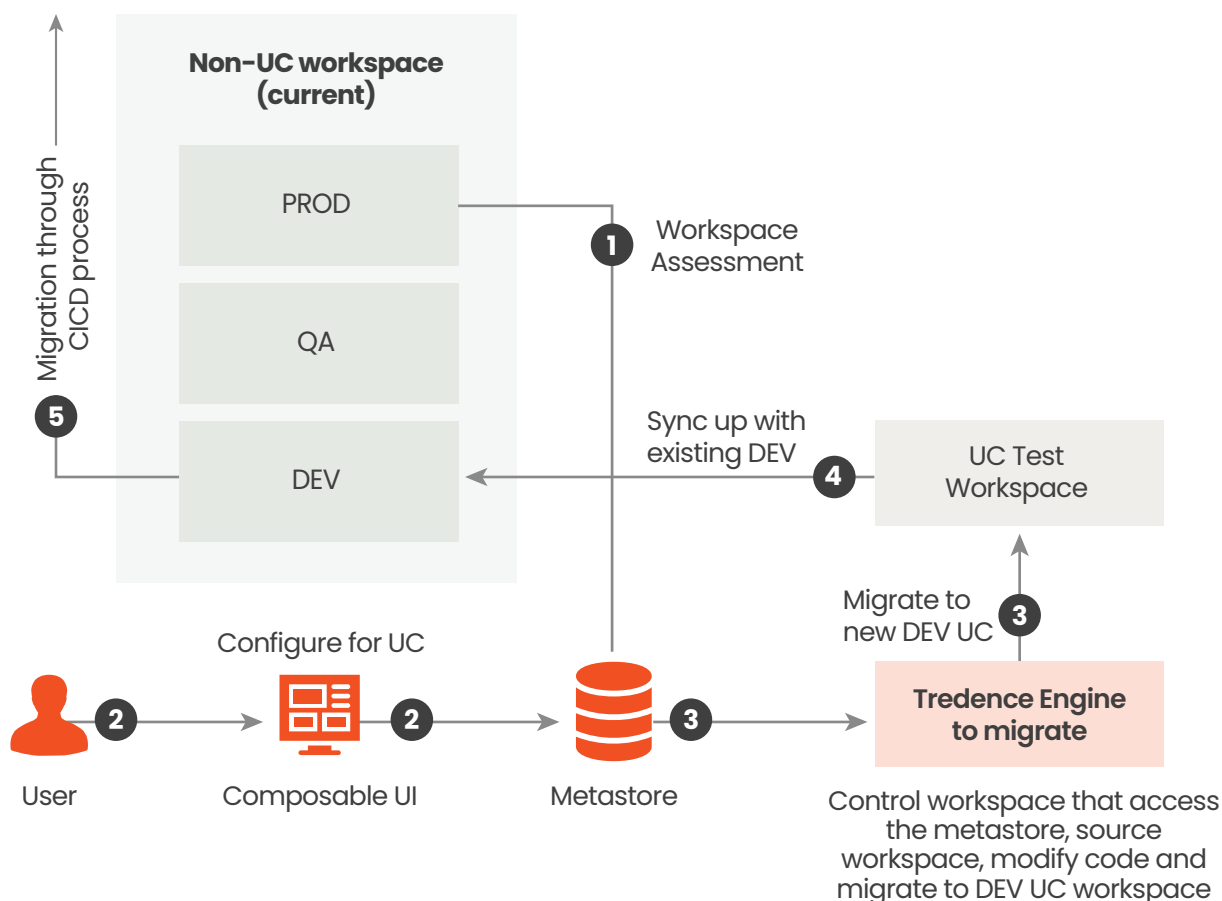
4. Retrofitting and migrating the code base

Next, UnityGO! solutions retrofit scripts to a sandbox Unity Catalog-enabled workspace, automates the creation of DDL scripts, notebooks, ACL scripts, and Databricks workflows. Teams can run scripts in a sandbox environment to migrate their objects to Unity Catalog and test without impacting existing applications. Then, they enable an existing application workspace for Unity Catalog and migrate the code from the sandbox environment to the existing development workspace, test and migrate it to quality and production workspaces through existing continuous integration and delivery (CI/CD) processes.

Review the UnityGO! Architecture







UnityGO! uses a migration approach that parallels development without impacting existing applications. As a result, teams can migrate existing production code.

Design – Solution Architecture



UnityGO! streamlines the migration of data at scale to the Unity Catalog.

Key Deliverables Teams Get with UnityGO!

KEY DELIVERABLES	
 Tables and Workflows migrated to UC	Migrate the external tables in Hive metastore to Unity Catalog based on configuration provided in the Tredence UI accelerator tool. Migrate and retrofit workflows for Unity Catalog
 Managed tables migrated to UC	Provide scripts to migrate managed tables in workspace to external tables and then catalog the metadata in Unity Catalog
 Code retro fit to read and write from UC	Code base (notebooks) to read and write data from Unity Catalog which are fully configured to be able to migrate across environments taking into account changes in the datalake and catalog environment
 Users and groups migrated to UC	Migrate existing users and groups that have table level and schema level ACL access to unity catalog with the same level of permission
 Replace local workspace groups with account level groups	Replace local workspace groups with account level groups at workspace resource level like clusters, cluster policy, folder access, SQL warehouses etc.
 Clusters enabled to run on UC	Migrate existing clusters to UC enabled clusters

Why Work with Databricks and Tredence

Databricks and Tredence take a solution-led approach to develop and deliver end-to-end enterprise data-driven solutions that solve last-mile challenges with data migration and modernization, analytics, and AI. Tredence provides purpose-built solution accelerators, built on Databricks Data Intelligence Platform, that solve domain and industry challenges at speed and scale and unlock millions of dollars in new revenues and cost savings.

Tredence is an elite partner of Databricks and hosts Brickbuilders solutions in every vertical served by Databricks.





Gain access to our combined solution

Databricks Data Intelligence Platform

provides one platform for integration, storage, processing, governance, sharing, analytics, and AI and is available via Amazon Web Services, Google Cloud Platform, and Microsoft Azure.

Databricks Unity Catalog

enables enterprises to seamlessly govern structured and unstructured data, ML models, notebooks, dashboards, and files on any cloud or platform.

UnityGO!

streamlines migration to the Databricks Unity Catalog with a simple, repeatable process. Co-developed with Databricks, it is the only approved automated migration tool for Unity Catalog.



Customers realize the value of Databricks – Tredence better together



Gain a universal data set catalog

Unity Catalog makes data sets securely accessible to internal teams and approved external partners while eliminating the risks of data redundancy, corruption, and exposure.



Access new features with Unity Catalog

Data Intelligence Platform IQ, Data Intelligence Platform Federation, and Delta Sharing are all enabled through Unity Catalog, enabling teams to gain richer insights and monetize organization data assets.



Leverage end-to-end services

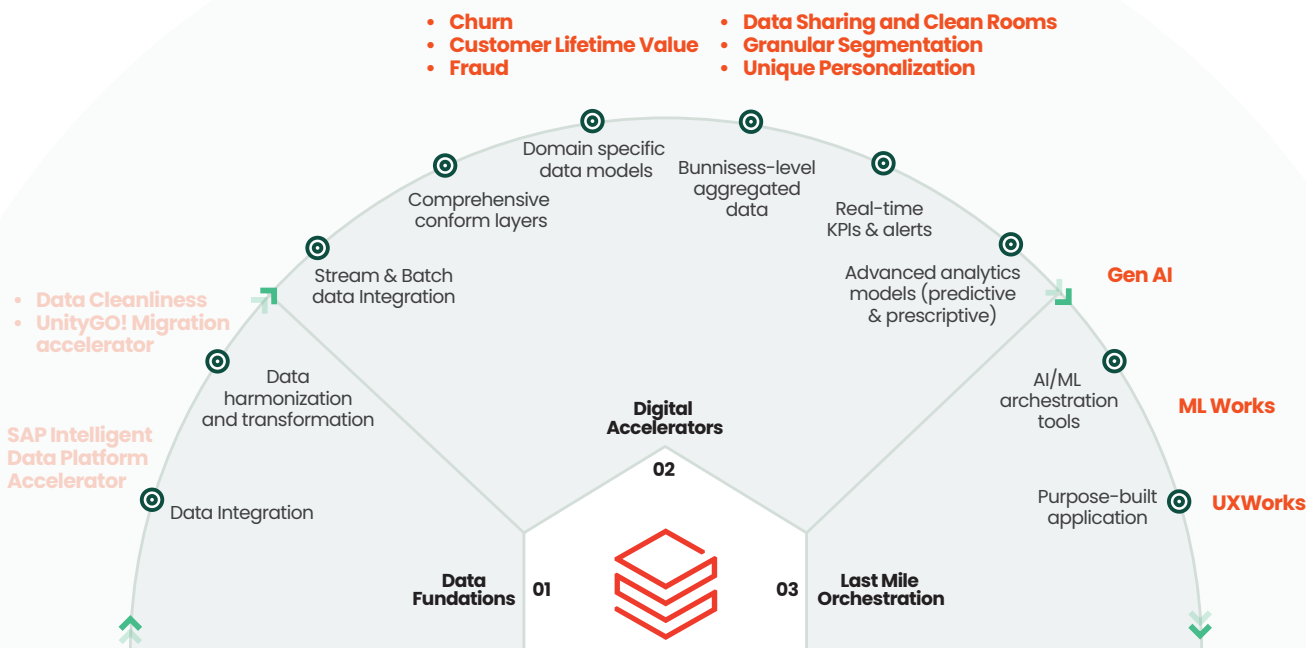
Tredence provides end-to-end services for data, analytics, and AI, including data migration and modernization; business intelligence and analytics enablement; LLM engineering, development, and operations; and platform engineering capabilities.



Work with proven partners

Databricks and Tredence collaborate on developing discriminative and generative AI solutions that solve industry challenges and provide speed to value. In addition, Tredence has extensive expertise in migrating and modernizing the data of Fortune 500 companies.

Harness Tredence Accelerators to Solve Complex Challenges



Tredence accelerators unlock more value throughout the data lifecycle while solving complex industry challenges, such as reducing customer churn and fraud.

Talk to Tredence. Get started with UnityGO!



Enterprise leaders want to democratize and monetize data to drive new revenues and cost savings. By partnering with Databricks and Tredence, you gain access to the skillsets, frameworks, and tools you need to unify data to enable collaboration and secure sharing.

Streamline the path to generating insights, develop a more robust picture of your customers and business opportunities, and accelerate time to value with the Databricks Unity Catalog and UnityGO!

Seize This No-Risk Offer

Ready to get started?

Contact us to schedule a 60-minute discovery call, where we'll learn about your needs, discuss your migration path, and share success stories of companies that have unified their data.

Migrating and modernizing your data is within reach. Create a data-driven culture and reap the financial rewards of metrics-driven decision-making.

[KNOW MORE →](#)



Access These Resources

Use these resources to learn more about the Databricks Unity Catalog, the benefits it provides, and the key processes it enables.

Databricks

-  [Unity Catalog](#)
-  [Unity Catalog best practices](#)
-  [A Practitioner's Guide to Unity Catalog—A Technical Deep Dive](#)

Tredence

-  [Introducing Unity Catalog: Governing Data with Databricks \(blog\)](#)
-  [What Is Unity Catalog and Why Is It a Game Changer? \(blog\)](#)
-  [Setting Up Unity Catalog and Object Organization \(blog\)](#)
-  [Understanding Databricks Cluster Types in Unity Catalog \(blog\)](#)
-  [Mastering Data Governance with Unity Catalog](#)
-  [Effective Migration to Unity Catalog: A Guide](#)

ABOUT THE AUTHOR



Maulik Divakar Dixit

Director and Sr. Azure,
Databricks and Cloud Architect
Tredence Inc.

Maulik has 20+ years of experience in designing and implementing large scale data and analytics solutions for Fortune 100 companies. He has been instrumental in building end-to-end data engineering solutions from data ingestion to data processing to serving data to customer.

Maulik has worked in the capacity of Azure and Databricks architect with several leading clients. Some of the transformational engagements he has spearheaded include enterprise data lake implementation on Databricks where he worked with the business to define and build supply chain data domains, and building a business data lake on Azure Databricks for a large CPG client.



David LeGrand

Senior Director, Partner Development
and Platform Evangelism,
Tredence Inc.

Execution focused Industry Go-to-Market leader delivering successes spanning marketing thought leadership campaigns, through post-sales value solution delivery. It has been enabled by an ecosystem of partners, ISVs and GSIs producing a legacy of digital transformation, risk mitigation and operational excellence growth.

+Twenty years of driving successful growth and portfolio development in the Aerospace, Chemical, Consumer Goods, Electronics, Energy, Industrial, Retail and Transportation markets through strong customer and partner interaction.

Want to know more about us?

Please visit: www.tredence.com

Follow us at: [in](#) [t](#) [v](#) [f](#)