TREDENCE

Beyond Possible

A Wholesale Food Giant's Data Modernization Journey Accelerates Analytics Success

CASESTUDY

ABOUT THE CLIENT

The client has grown from a family business to a multinational delivering culinary solutions to food service businesses and supermarket chains in over 112 countries. One of North America's largest private companies with over $4 billion in revenue, it has a portfolio of over 2000 innovative products that help business owners delight customers' palates with global flavors that have a local touch.

One of the primary values underpinning this success is collaboration. The company believes in working together to meet the changing needs of all its stakeholders. For this, the company wanted all its data to sit on an enterprise-grade advanced platform that offered a unified truth to all business functions, frameworks that avoided reinventing the wheel, and complete preparedness for agile decision-making driven by state-of-the-art analytics.

The Need to Unify Siloed, Unprepared Data for an Insights-led Digital Future

As a large, successful organization, the client had captured massive volumes of data over the years, tracking its daily operations and external and internal initiatives. With the organization now poised for a data-driven future, decision-makers realized that this data had concerns that would hamper competitive agility.

It sat on eight to ten file system types, including worksheets, Oracle databases, and flat files, which meant even BAU analysis within a business function lost time to employees toggling between screens, consolidating and preparing data.

Since there were no reliable ways of recording data cleansing and consolidation processes, there was duplication of efforts within the same team or across teams.

In addition to being scattered, the data was in diverse formats and did not always meet data quality requirements. Therefore, harnessing it for insights and KPIs was sometimes not possible and at other times, off the mark.

The data did not lend itself to complex, strategic analytics initiatives that are increasingly pivotal to modern organizational success.

The combination of these issues and rapidly growing data volumes made cross-functional collaboration increasingly infeasible.

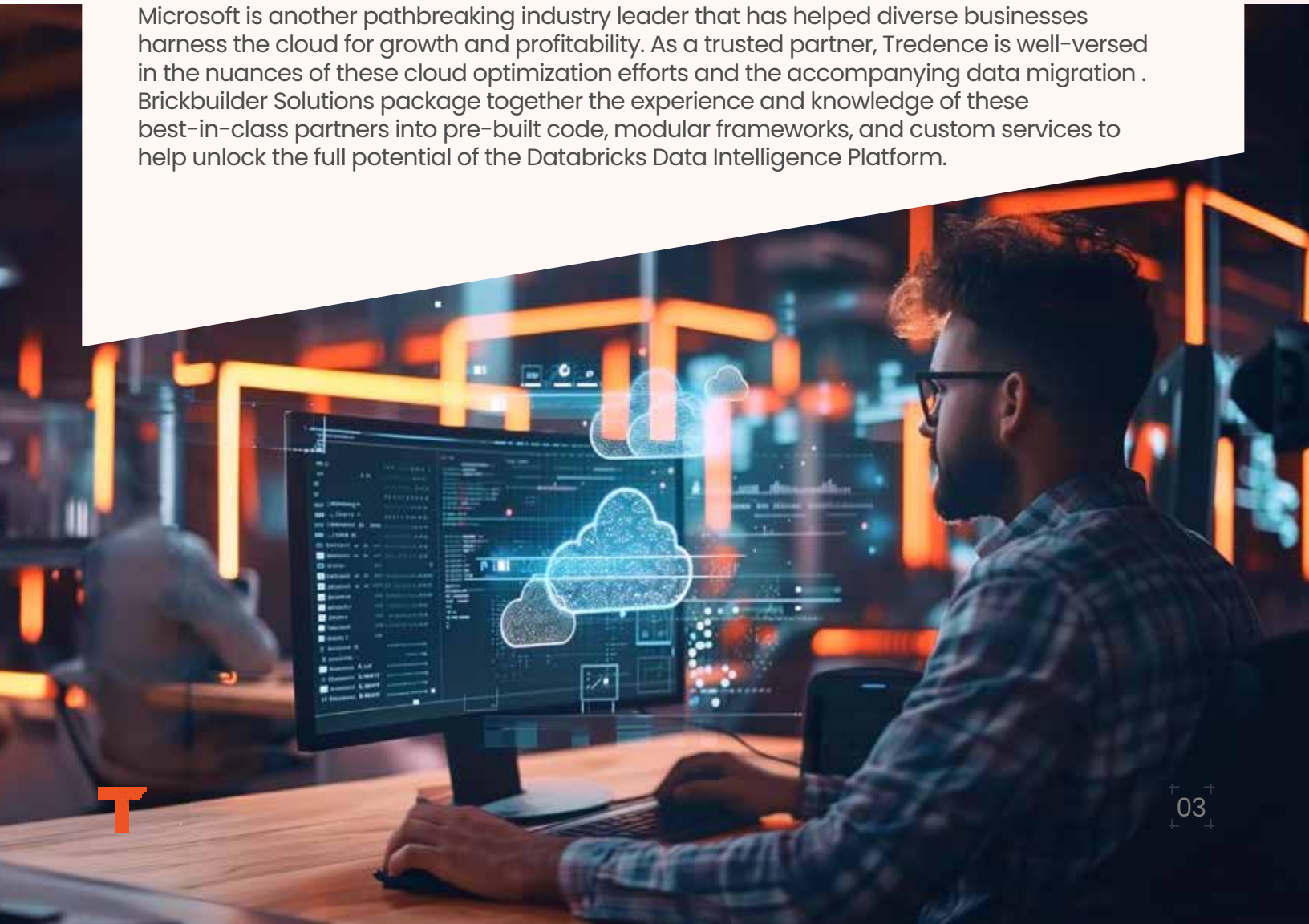
It was time to set up a cloud-based, unified, enterprise-grade data platform with cutting-edge data engineering capabilities to drive collaborative insights. As a mammoth project to provide a data foundation to propel an exponential digital journey, it needed a multi-capability network of partners that would come together seamlessly. The client saw that Tredence has solid credentials in data engineering and a successful track record of working with industry-leading partners Databricks and Azure.

The Tredence Partner Ecosystem Bridges the Cloud Data Gap for Businesses

With companies moving to the cloud, migrating the data environment seamlessly and securely is crucial. Over the last few years, Tredence's Brickbuilder Solutions team has executed many large-scale data migrations for multinationals across sectors in partnership with data intelligence leader Databricks and pioneering cloud provider Azure.

Databricks has a range of cutting-edge, innovative data platforms. One of these is the Lakehouse architecture, which marries the flexibility of a data lake with the structural robustness of a data warehouse. Tredence has helped deploy all these platforms, including Lakehouse, to unlock the power of data for numerous clients.

Microsoft is another pathbreaking industry leader that has helped diverse businesses harness the cloud for growth and profitability. As a trusted partner, Tredence is well-versed in the nuances of these cloud optimization efforts and the accompanying data migration . Brickbuilder Solutions package together the experience and knowledge of these best-in-class partners into pre-built code, modular frameworks, and custom services to help unlock the full potential of the Databricks Data Intelligence Platform.



How Tredence's Brickbuilder Solutions Drive Analytics Value with Databricks Lakehouse and The Unity Catalog

To understand how 2000 products appeal to palates around the world, how to manage relationships with supplier networks and customers worldwide who are food service businesses and retail companies, and how to empower and engage with employees, the client needed the chosen architecture to provide huge flexibility in the types of data ingested. Flowing in from conventional sources and non-mainstream ones like social media platforms, this data had to comply with privacy and governance norms and be ready for advanced analytics and AI/ML applications spanning the business.

The blended nature of the Databricks Lakehouse platform delivered these criteria. The platform also creates a single source of truth that effectively tackles the issue of hampered collaboration. Adding another layer of uniformity, the Unity Catalog ensures that data governance is standardized across the company.

The Tredence Brickbuilder Solutions team was asked to create and implement the complex, pan-organizational Databricks Lakehouse platform on Azure. With its intersectional expertise and strong partnership experience, the team was seen as the best fit to customize the data platform to the business goals and vision, user needs, and the long-term technological viability crucial for the RoI of such a massive investment .

The Food Giant Builds a Best-in-Breed Data Foundation for Analytics Success

The Tredence Brickbuilder Solutions team sat down with the client to understand the manifold objectives and simplified them to a few key implementation goals.

The food giant's enterprise-grade data lakehouse platform on Azurecloud would have to:

Modernize data

Streamline ingestion of diverse data. Automate data quality capabilities, including alerting and monitoring, to ensure accuracy, consistency, and compliance

Prepare it for strategic analytics

Institute a centralized and connected data model so information within and across teams could be brought together rapidly at scale to light up intra-team as well as cross-functional use cases with accurate measuring of outcomes.

Over one and a half years, a multi-location Brickbuilder Solutions team

- Stood up the pan-organization data platform using Databricks Lakehouse and the Unity Catalog
- Leveraged and customized the T-Ingestor solution to meet client-specific requirements of enabling ingestion of structured and unstructured data through Yaml files
- Used the T-Assurer framework to test the ETL pipelines in an automated manner
- Leveraged proven data quality frameworks supported by robust logging and monitoring and easy-to-use dashboards to ensure accurate, compliant data from the Bronze layer onwards
- Used the encryption framework on Databricks to encrypt HR, customer, and vendor data from source systems and store them as encrypted values in Bronze, Silver, and Gold layers of the Data Lakehouse
- Built the Master and Reference data layers with the ability to harmonize data across multiple domains
- Empowered users across domains such as Marketing, Supply Chain, and HR with a single collaborative source through a centralized Data Model that lights up more than 100 KPIs. Each domain was entrusted with ownership over its data.

Bronze, Silver, and Gold layer	T-Ingestor	T-Assurer	Yaml files
The bronze (raw), silver (validated), and gold (enriched) denote the quality of the data in each of these layers.	Databricks-powered codeless automated data migration accelerator with built-in quality, catalog, and metadata management features.	A home-grown framework of standardized practices, methodology, and accelerators to reduce the overall time to value	A semi-structured data file with enhanced flexibility for customization. Greater bandwidth for adding parameters, new tables, and frameworks for data-related functionalities.

Solution Framework:

- Metadata Driven
- Reusable Rules Engine
- Pre-Configured Rules (14)
- Detailed Data Quality Report
- Automated Email on Quality
- Extensible to add new rules
- Enables using Custom SQL
- DQ Dashboards

Solution Components:

Metadata Tables

- Set of 2 metadata tables in SQL DB
- One for Rule configuration per object per column on which the rule needs to be applied
- Second for Object configuration with business email id and object level settings

Rules Engine

- Azure Databricks Notebook which reads the metadata table
- Dynamic query generation to apply rules as configured in the metadata table for the object and column
- Load the records with DQ issue in a separate table with reason code into SQL DB

Email and Reporting

- Azure Databricks notebook to read the Object configuration table
- Load the bad records file into storage account, SQL DB and call Azure Logicapp using REST API
- Send email with DQ summary and DQ summary, by Violations link to file
- Create a Power BI report for error reporting

Modules:

01 CORE ENGINE

- Core Engine that will take the DQ rules and generate a comprehensive DQ report sent by Email

02 DQ DASHBOARDS

- Build historical snapshots of Data quality results for each object
- Build Dashboards at application level, object level that trend of how data quality progress over time

03 METADATA AND GUI

- Ability to populate the RDBMS table with object and column metadata.
- Build GUI to enable users create rule in RDBMS

04 INTELLIGENT RULE GENERATOR

- Scan the representative sample files, understand the datatypes of column and automatically build DQ rules
- Users have the ability to select or reject rules created as part of intelligent rule creation

The Impact: A Future-Proof Platform Empowers Teams To Unlock Rapid Actionable Insights

3X

Improvement in time to value for ingesting data from similar source types

50%

Reduction in time to develop and market the use cases

50-60%

Improvement in the accuracy of descriptive and prescriptive KPIs

Once the implementation was done, client teams saw a significant, quantifiable improvement in how they engaged with data. They could leapfrog into developing use cases and tracking KPIs without having to prepare and consolidate them for analysis. Access to what other teams do has prevented duplication and doubled accuracy and efficiency.

Since the technology partners spearhead innovation in their respective spheres, the platform will remain robust and usable for decades to come. Hence, the client is firmly on the path to analytics success. Here is a look at some of the collaborative breakthroughs client teams achieved by leveraging the strengths and capabilities of the custom Brickbuilder Solutions Lakehouse platform.

Shattering Global Data Siloes for Impactful Digital Marketing

The client was investing strongly in digital marketing efforts to drive demand. However, these efforts were siloed within each country. Therefore, a common global metric was absent, and business leaders were unable to gauge the impact of spends on revenue or future demand creation, especially at a strategic level.

Tredence was tasked with creating a model to harmonize metrics across countries, quantify the impact of digital investments on demand, and work with multiple global teams to socialize this insights platform.

The Tredence team set up automated data pipelines to bring multiple global data sets from social media platforms into the data lake at specified time intervals. Harmonization and taxonomy using NLP tackled common campaign tagging and naming across channels and countries. An ML model was set up to decompose sales into base vs. offline marketing vs. digital marketing impact. Multiple stakeholder workshops ensured adoption and viability.

Time to global insights is set to drop by half and the impact of digital marketing to double within two years.

Multi-functional, AI-Powered Customer Acquisition in the C&U Market

The client wanted an analytical solution to identify potential opportunities within the College & University (C&U) channel. This would help identify specific core categories where existing customers were not engaging and potential headroom to grow through seeding these market gaps.

Five Tredenceans with cross-functional expertise worked with the client's IT, Supply Chain, and Finance teams to build a custom model. It ingested and unified internal data from the CRM, operator reports, transactions, claims, the customer and product master, and IoT data for shake dispensers. External data that comprised customer and location attributes was also used.

Algorithms were created for a cosine-based Customer Similarity Score and a Recommendation System for volume and probability. ML workflows that leveraged the Unity Catalog implemented the models. Their accuracy is being monitored for over a year, so a highly fine-tuned solution is handed over to the client.

About Tredence Inc.

Tredence is a global data science solutions provider focused on solving the last mile problem in AI. The 'last mile' is the gap between insight creation and value realization. Tredence is a Great Place to Work-Certified and as a 'Leader' in the Forrester Wave: Customer Analytics Services.

Tredence is 2500+ plus employees strong with offices in San Jose, FosterCity, Chicago, London, Toronto, and Bangalore, with the largest companies in retail, CPG, hi-tech, telecom, healthcare, travel, and industrials as clients.

Want to know more about us?
Please visit: www.tredence.com

Follow us at: