DataKnow

Methodology, analytical profiles and success stories

AI Midsize Summit 2025

DataKnow

About Us

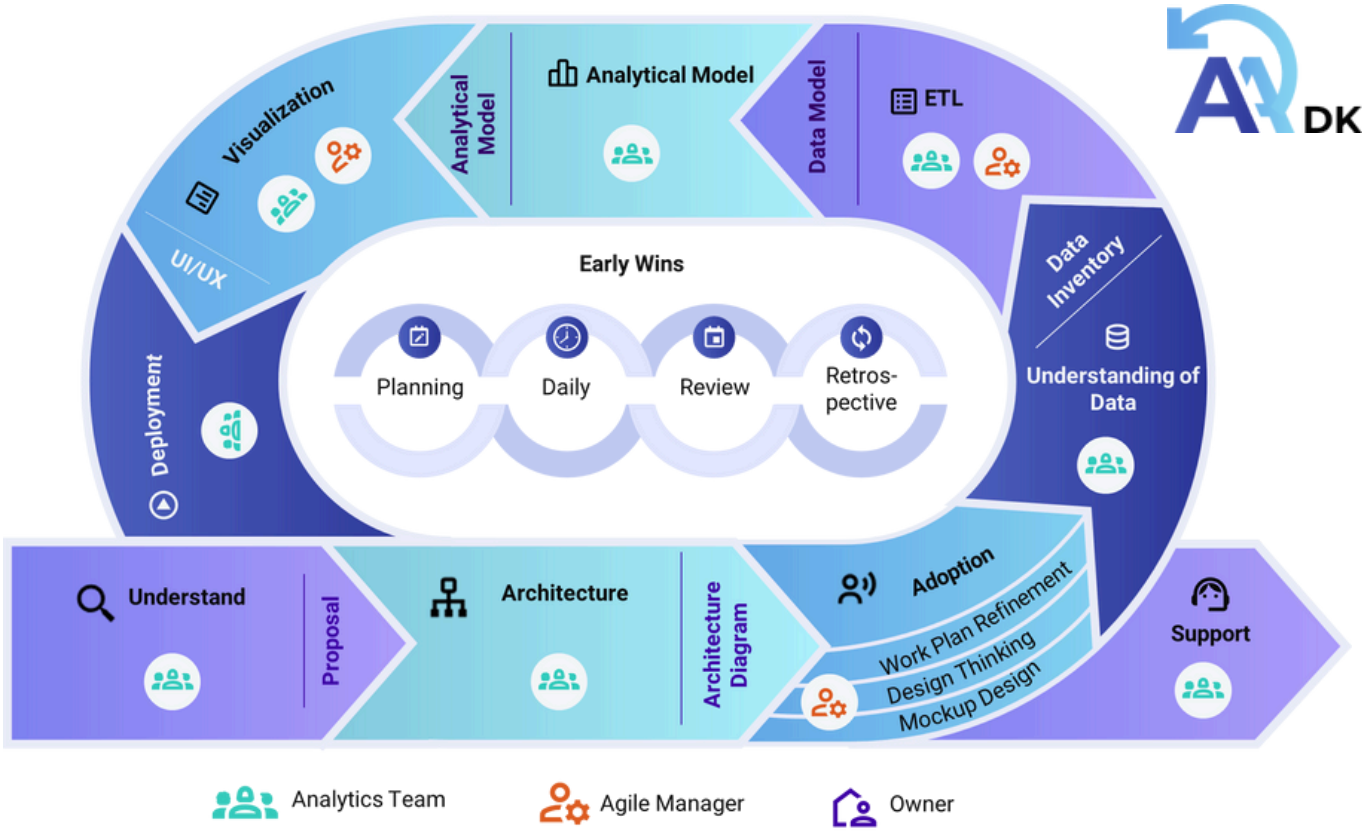
In DataKnow we believe in the power of data to transform the world. Our purpose is to democratize data analytics and artificial intelligence.

We are a leading data and analytics firm specializing in AI solutions. As a trusted Microsoft partner, we leverage advanced AI technology to provide transformative insights for businesses. We empower clients across industries to make data-driven decisions with confidence and efficiency. Partner with us to unlock the full potential of your data and stay ahead in today's dynamic business landscape.

More than 100 Consultants, Data Scientists and Data Engineers.

DataKnow

Agile Analytics Methodology



At DataKnow, we use Agile Analytics to find value in a dataset more quickly and to adapt better and more effectively to our client's needs.

Accelerated Value

Deliver project value faster and more efficiently.

Optimized Time

Boost productivity by making the most of work time.

Strategic Balance

Achieve flexibility and alignment between scope, cost, and timeline.

DataKnow

DataKnow Analytical Profiles

Data Scientist

Main Activities:

- Full vision of the process, starting from business understanding, identifying data to be used, models to be developed, deployment and monitoring.
- Exploratory and descriptive analysis, pattern discovery, and insights.
- Design and development of advanced analytics models, machine learning, and artificial intelligence.

Skills:

- Analytical mindset with high skills and abilities in both technology and business.
- All the skills of a data analyst, but with deeper business experience and data handling.
- Teamwork in collaborative and agile environments (Scrum certification) and strong communication skills.
- Broad knowledge in mathematics, statistics, programming, and visualization.
- High experience with algorithms in Analytics, Deep Learning, Machine Learning, and Artificial Intelligence.
- Strong programming skills in Python, R, Spark, SQL, and AWS components (Sagemaker, Kanbas, Studiolab, Databricks, EMR, etc.).

Deliverables:

- Insight and exploratory/descriptive analysis reports.
- Data quality reports and technical/functional documentation of models and results.
- Coding and programming of implemented models and pipelines.
- Visualization of business insights.
- Analytical models and results in business terms.

DataKnow

DataKnow Analytical Profiles

Data Architect

Main Activities:

- Define data structures including design, creation, supervision, and administration.
- Define data logic models and standards for data processing and governance.
- Adjust operations and ensure sustainable scalability of the technological infrastructure as the organization grows.

Skills:

- Curiosity about the state-of-the-art in technology implementations and future deployments and automations.
- Understanding of best practices and methodologies for data integration and agile frameworks (certification in Scrum).
- Experience with scalable and optimized cloud architectures (Cloud Architecture) AWS and Azure.
- Experience and knowledge in technological infrastructures (Docker, REST APIs, batch, ML APIs, etc.).
- Experience and knowledge in Data Governance, data security, operationalization of models (DevOps, MLOps), and code versioning (Git, AWS CodePipeline, etc.).

Deliverables:

- Multi-layer ETL, data load logic, mapping, workflows, and stored procedures.
- Data Flow Diagram (DFD). Data model design, As-is & To-Be process.
- Service-Oriented Architecture (SOA), ensuring best performance.
- Governance-oriented data models.

DataKnow

DataKnow Analytical Profiles

Agile Manager

Main Activities:

- Regular follow-up meetings with the team.
- Coordinate project stakeholders and consolidate the team.
- Lead the team to achieve project objectives.

Skills:

- Project management skills for transforming ideas into analytical products.
- Agile collaboration and communication (Scrum certification).
- Focused on problem-solving and innovation ideation.
- Interest in tech trends and deployments.
- Tactical and operational project coordination.
- Organizational, planning, time-tracking, and project closure capabilities.
- Budget planning and administration.
- Specialized leadership in advanced analytics implementation.
- Scope and planning management for analytics projects.
- Best practices knowledge for data integration and agile frameworks.
- Experience with technology infrastructures.
- Background in administration, statistics, economics, systems/industrial engineering, finance.

Deliverables:

- Project management, budget, and progress reports.
- Key success stories, metrics, and business impacts.
- Strategic planning and team structuring by profile.
- Manage project constraints (scope, timeline, cost, budget, quality).

DataKnow

DataKnow Analytical Profiles

Data Engineer

Main Activities:

- Build and automate scripts for extraction, transformation, and loading (ETL).
- Program and organize the company's data pipelines and flows using standard data languages (SQL, NoSQL, Spark, Glue, Pandas, Hadoop).
- Define and configure continuous integration specifications. Perform data analysis using OLAP techniques.

Skills:

- Goal-oriented, autonomy for self-management, collaborative and agile work.
- Curiosity in technical implementations and open-mindedness to multiple data languages.
- Ability to handle high volume and dimensionality.
- Experience building Data Warehouses from analysis, design, ETL, reporting, and dashboarding (Power BI, Qlik, Tableau, AWS QuickSight).
- Experience in developing corporate applications (Power Apps, Power Automate, .Net).
- Knowledge of distributed data storage systems (HDFS), distributed processing (Hadoop, Spark), and OLAP techniques.
- High proficiency in SQL, NoSQL, Spark, Python, Pandas, Hadoop, and AWS components (Redshift, Glue, Sagemaker, Athena, DynamoDB, Databricks).
- Experience in information security, code versioning (Git), and DevOps.

Deliverables:

- Technical documentation on pipeline configuration.
- Data sheet dictionaries and data quality reports.
- Data models, Analytical Base Table (ABT).
- Interactive dashboards visualizing business insights.

DataKnow

DataKnow Analytical Profiles

Data Governance

Main Activities:

- Define data governance policies, review relevant laws and regulations.
- Monitor and report on data quality metrics.
- Consult on best practices in data governance.
- Skills:

Skills:

- Experience managing governance projects (planning, execution, monitoring).
- Familiarity with governance frameworks (DAMA, DGMM, DCAM).
- Implementation experience of governance policies and processes.
- Knowledge of data security and personal data protection laws.

Deliverables:

- Detailed report assessing current data governance maturity.
- Strategic plan outlining vision, goals, and tactics for governance.
- Document with policies, norms, and procedures.
- Interactive dashboard for real-time data quality monitoring using KPIs.

DataKnow

DataKnow Analytical Profiles

Data Artist

Main Activities:

- Be creative, with high innovative development, capturing client needs through Design Thinking techniques for dashboards and data visualizations.
- Think "outside the box", use intuition, and communicate data behavior in a simple and interpretable way.
- Design and create mockups of dashboards and reports.

Skills:

- Proactivity in problem-solving and attention to detail.
- Critical thinking for information management.
- Communication skills for summarizing large data volumes clearly and precisely.
- Teamwork in agile settings (Scrum certification) and Design Thinking.
- Basic knowledge of tools like Power BI, Tableau, QuickSight, Qlik.

Deliverables:

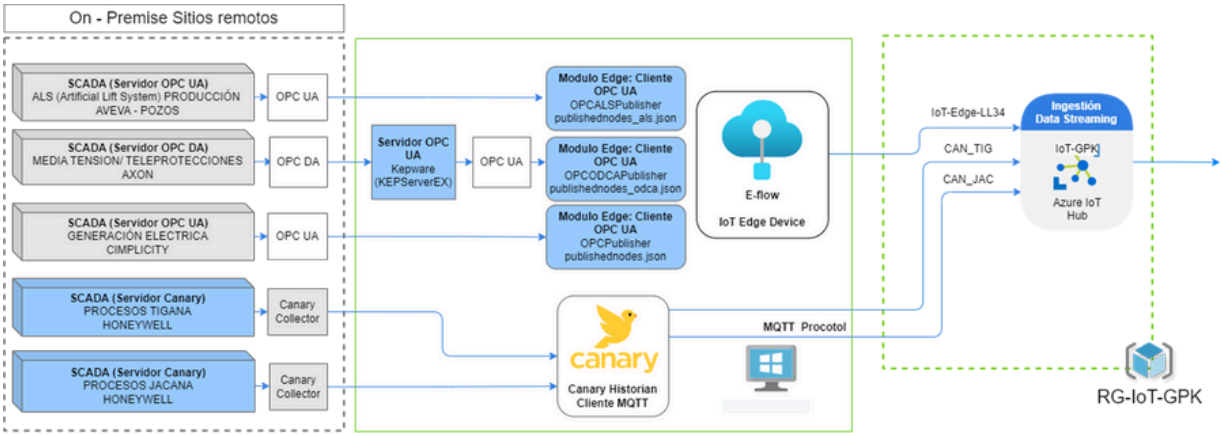
- Mockups for business visualizations (reporting), dashboard development.
- Data visualizations and infographics conveying easy-to-understand messages.

DataKnow

Experience

Success Case

IoT – Real-Time Telemetryfrom
SCADA Systems in the OT Network



What did they need?

- To monitor real-time data from various production processes sourced from different SCADA systems, send it to the cloud for monitoring (dashboards), and once the data is available, perform analytics on it.

What did they achieve?

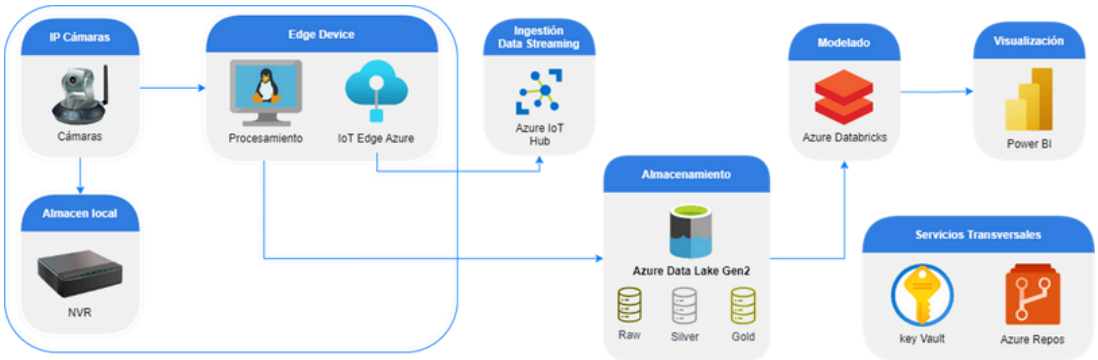
- Connection of SCADA systems using the OPC UA protocol to IoT Edge E-FLOW deployed on a virtual machine within the OT network.
- Connection of SCADA systems to a local historian and transmission of data using the MQTT protocol to IoT Hub in the cloud.
- Telemetry transmission from IoT Edge to IoT Hub.
- Real-time telemetry visualization with business analytics applied via Data Explorer dashboards.
- Data storage in real-time: Using Stream Analytics, data is sent to a Storage Account, organized by source, and subjected to transformations for proper storage.
- Business transformations applied to the data for integration with other monitoring processes and software.

DataKnow

Experience

Success Case

Computer Vision



What did they achieve?

Limestone Detection Model POC:

- Size classification model (small, medium, large and very large) of limestone transported by cement belt, for the percentage calculation of each of the stone sizes in a given time.

Detection and classification model of products on shelves

- Model that detected the products on a shelf and classified them by product, thus obtaining relevant metrics

Model for analysis of inputs, outputs and times by PoC zones

- Model that detected people through CCTV cameras to control entries and events where each employee had a number that the model identified for tracking through the site.

DataKnow

Experience

Success Case

ML – Retail Sector – Demand Forecasting for Product Sales

What did they need?

- A demand forecasting tool to predict future sales for the TAT distribution channel on a national scale.
- A solution to manage inventory levels effectively based on demand predictions.
- Improved accuracy in forecasts to enhance operational and financial KPIs.
- Better cash flow management through demand-driven insights.

What did they achieve?

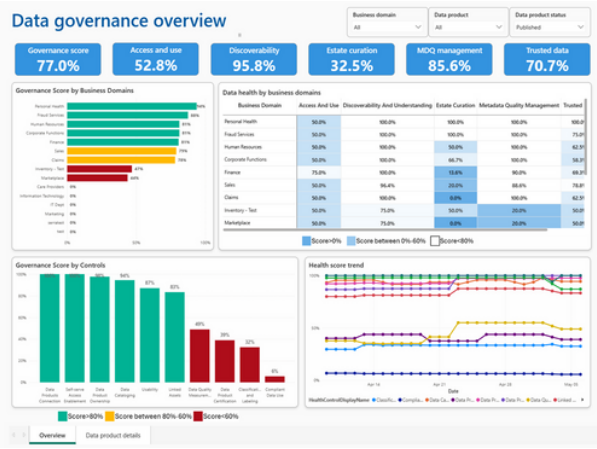
- Accurate sales projections calculated using machine learning and deep learning techniques in the cloud.
- Improved inventory management to prevent stockouts or unnecessary overstocking.
- Synchronization of sales, planning, and production areas for better operational alignment.
- Significant MAPE improvement, reducing it from over 20% to a range between 12% and 6%.

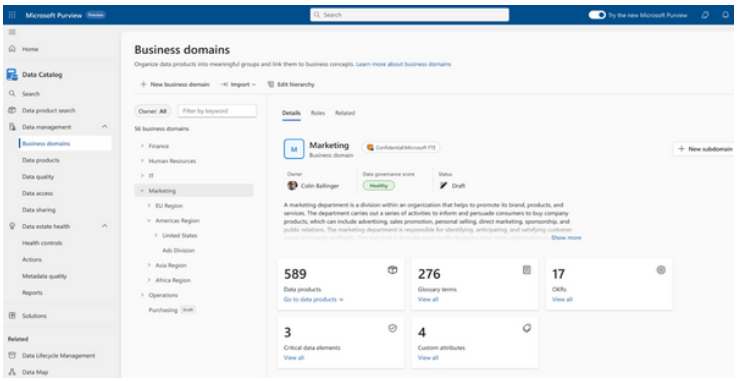
DataKnow

Experience

Success Case

Purview in financial





What did they need?

- The company is in the process of migrating its data warehouse to the cloud, requiring a tool to manage the organization's defined data governance. The goal is to implement Microsoft Purview, which enables unified governance and bridges the gap between technical and business language, contributing to data democratization.

What did they achieve?

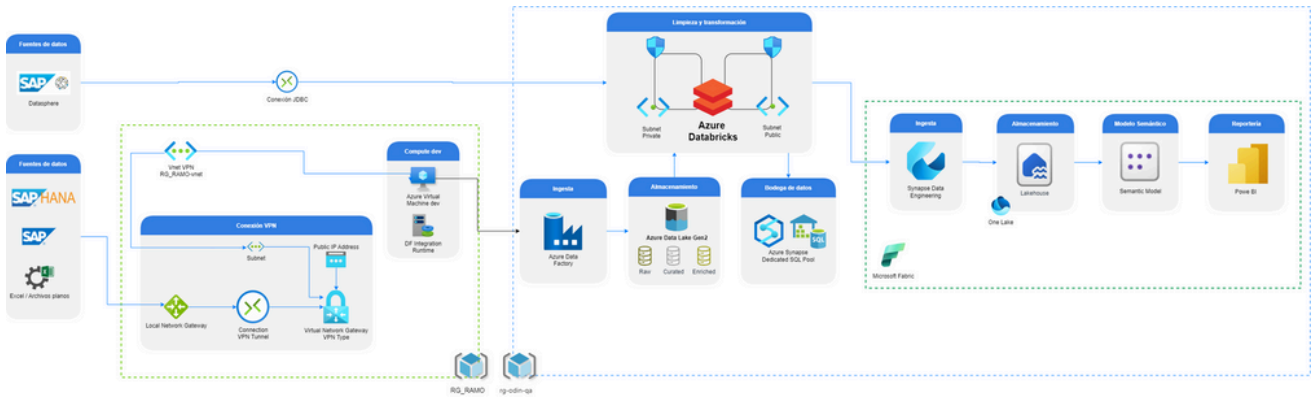
- Comprehensive Data Governance with Microsoft Purview: Developed a unified approach to data management, including inventorying assets, defining roles, and creating a holistic data map.
- Business Alignment: Defined business domains, key concepts, and a glossary to bridge technical and business understanding.
- Governance Reporting: Configured reports for domains, data products, and assets to monitor compliance and performance.
- API Integration and Training: Integrated REST APIs for extended capabilities and trained teams to effectively utilize Microsoft Purview.

DataKnow

Experience

Success Case

Microsoft Fabric + Databricks



What did they need?

- Create a scalable and flexible environment capable of extracting data from SAP Datasphere, SAP HANA, SAP and Sharepoint
- Optimize expenses related to cloud infrastructure without compromising the integrity and efficiency of existing data flows.
- Use Power BI for empowering the organization with actionable insights for informed decision-making.

What did they achieve?

- Catalog management and continuous monitoring of internal Azure Databricks data, enabling efficient administration.
- Successful migration of the "Cubo Comercial," "Cubo Abastecimiento," and "Cubo Financiero" models from Analysis Services to the Microsoft Fabric environment (approximately 3.5 billion records).
- Improved loading times and performance of final reports.
- Optimization of Databricks processing logic, achieving a 20% reduction in execution time

DataKnow

Contact us

www.dataknow.co

Leon Felipe Álvarez

+(57) 318 644 8190

leon.alvarez@dataknow.co

/in/leonfelipealvarez/

Let’s connect on LinkedIn