SIGMOID

Data Warehouse Modernization on Azure

www.sigmoid.com


About Sigmoid

Sigmoid is an emerging leader in data engineering and AI solutions.

750+
Employees

Work with **30+**
Fortune 500 firms

>97%
CSAT score

200+
ML models operationalized

5000+
Data pipelines built

Backed by
SEQUOIA

Awards and Recognition

500 Technology Fast 500 2023 NORTH AMERICA Deloitte.

ISG Provider Lens 2023 Quadrant
Analytics Services
Rising Star, U.S.
Report releasing Jan 2024

Inc. 5000 America's Fastest-Growing Private Companies

DATA BREAKTHROUGH AWARD 2021
Open Source data solution provider of the year

FORRESTER
Now Tech: AI Consultancies, Q1, 2021 Report

Major Contender in
EVEREST GROUP
Analytics and AI Services Specialists
PEAK Matrix (2022)

Offices


New York


San Francisco


Dallas


Lima


Bangalore


Amsterdam


London


Sao Paulo

SIGMOID

Enabling Business Transformation with Full-Service Capability Suite

Business Consulting & Data

- Data Strategy & Vision
- Data Monetization
- Data & Technology Roadmap
- Technology Evaluation & Selection
- Data Governance & Security Strategy
- AI/Gen AI Strategy

Data Engineering Services

Data Pipelines	ML Engineering	Cloud Trans.	BI / Consumption
Data Migration & Conversion	Model scaling & productionizing	Cloud Migration	Data Lake / Mesh
Performance Optimization	Feature Engineering	Application Modernization	Data Product
Data Ingestion ETL/ELT	Pipeline Optimization	Cost optimization	BI Reporting & Visualization
			AI/ML, LLM

Data Science

- Supply Chain Analytics
- Marketing & Consumer Analytics
- Operational Analytics
- E-Commerce & Sales Analytics

Managed Services

- Data Labs
- Cloud Infra Support and Management
- Devops and Secops Support
- DataOps & ML Ops
- Data Application Managed Services

Governance & Security Services

- Data Catalog & Lineage
- Master Data Management
- Data Quality & Security

Technology Expertise

Technology Partners

Microsoft, databricks

Cloud Technologies

python, PySpark, jupyter, TensorFlow, R, spaCy NLP, MATILLION, DATADOG, Spark, hadoop, cassandra, mongoDB, HBASE, atlan, Alation, kubernetes

SIGMOID

Sigmoid Capabilities - Experience in implementing data solutions in Azure

Sigmoid has worked with more than Five large customers to design, build and deploy solutions in Azure

Data Processing & Transformation:

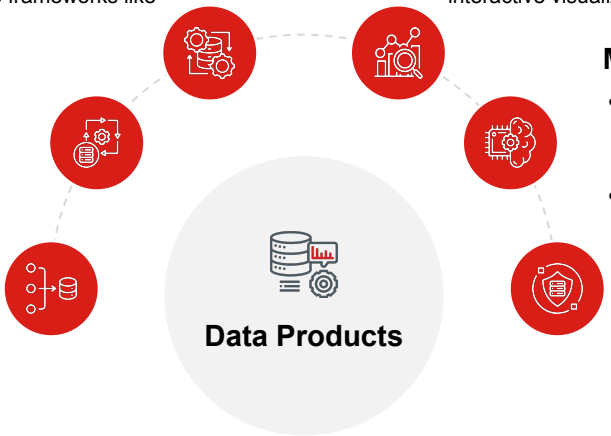
- **Azure Databricks:** Collaborative Apache Spark-based analytics platform to be used for big data processing and machine learning.
- **Azure HDInsight:** Managed cloud service for processing big data using popular open-source frameworks like Hadoop and Spark.

Data Analytics & Visualization:

- **Azure Synapse Analytics:** Analytics service which will be used for analyzing large amounts of data using either serverless or provisioned resources.
- **Power BI:** Business intelligence tool to be used for creating interactive visualizations and reports.

Data Storage & Management:

- **Azure Data Lake Storage:** Scalable and secure data lake for storing large amounts of structured and unstructured data would be considered.
- **Azure SQL Database:** Managed relational database service for structured data storage.



Machine Learning & AI:

- **Azure Machine Learning:** End-to-end platform for building, training, and deploying machine learning models.
- **Cognitive Services:** Pre-built AI services for vision, speech, language, and decision-making.

Data Ingestion & Integration:

- **Azure Data Factory:** Creating data pipelines to move and transform data from various sources.
- **Azure Event Hubs:** Real-time data ingestion from applications, devices, or any data streams would be done.

Security & Compliance:

- **Azure Active Directory:** Identity and access management service.
- **Azure Policy & Blueprints:** Tools for implementing governance and compliance across Azure resources.

Sigmoid's implementation of solutions in Azure involves leveraging a combination of services and tools tailored to specific business needs. Sigmoid would collaborate between data engineers, data scientists, business analysts, and other stakeholders as it is essential to align the implementation with business goals and ensure success.

Microsoft Solutions Partner

SIGMOID

Problems We Solve for Our Clients in the Data Landscape

Data Pipelines

- Business not able to respond in real time
- Automating data ingestion & scalability of data pipelines
- Data quality issues and trust in data
- Maintaining and monitoring data pipelines

MLOps

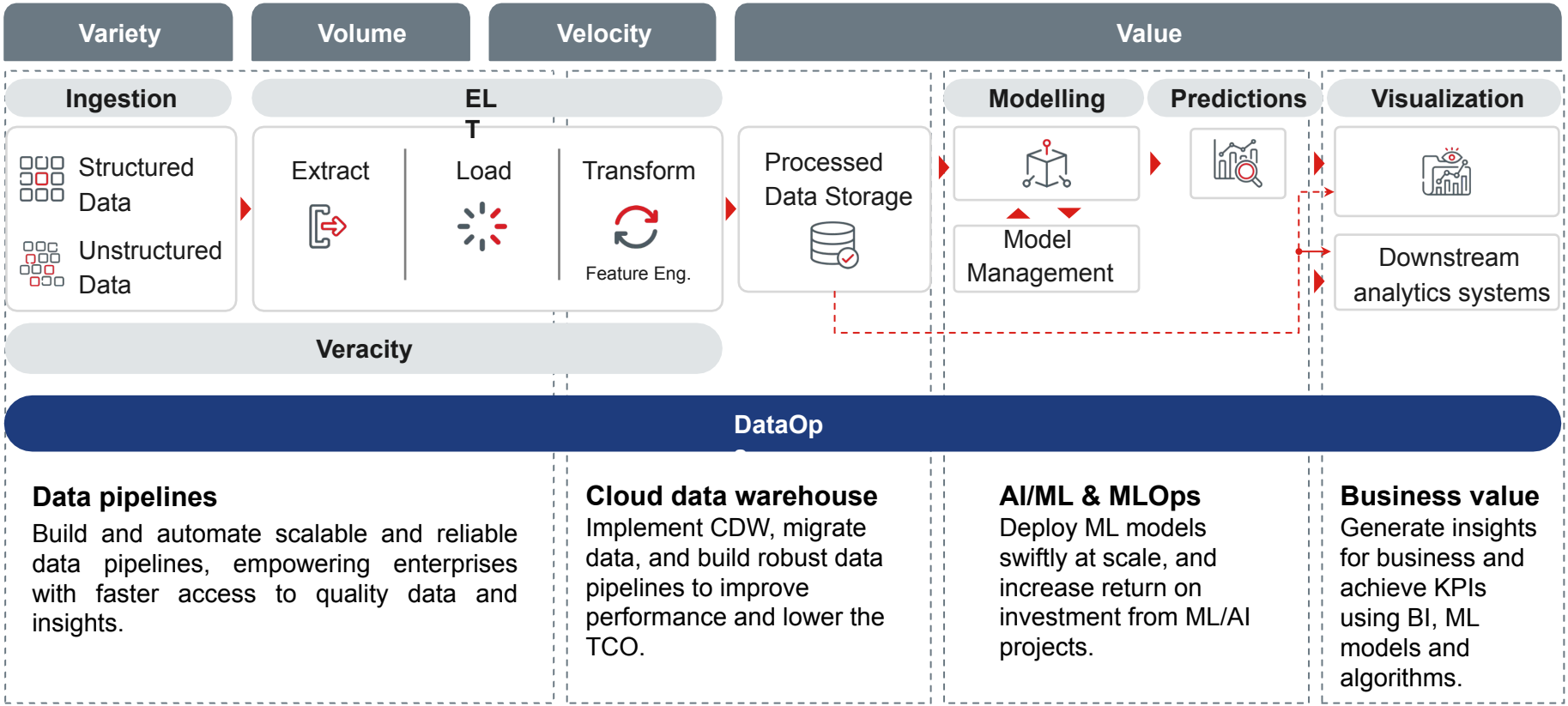
- High failure rate of deploying ML models in production
- Creating and managing ML pipelines
- Model Drift in Machine Learning
- Longer development and deployment lifecycle

Cloud Data Warehouse

- Technology selection for a use case
- Data migration from on-premise to CDW
- High TCO of a traditional data warehouse
- Rigid and inflexible infrastructure

SIGMOID

Services Spanning the Data Engineering and Analytics Value Chain



Data pipelines
Build and automate scalable and reliable data pipelines, empowering enterprises with faster access to quality data and insights.

Cloud data warehouse
Implement CDW, migrate data, and build robust data pipelines to improve performance and lower the TCO.

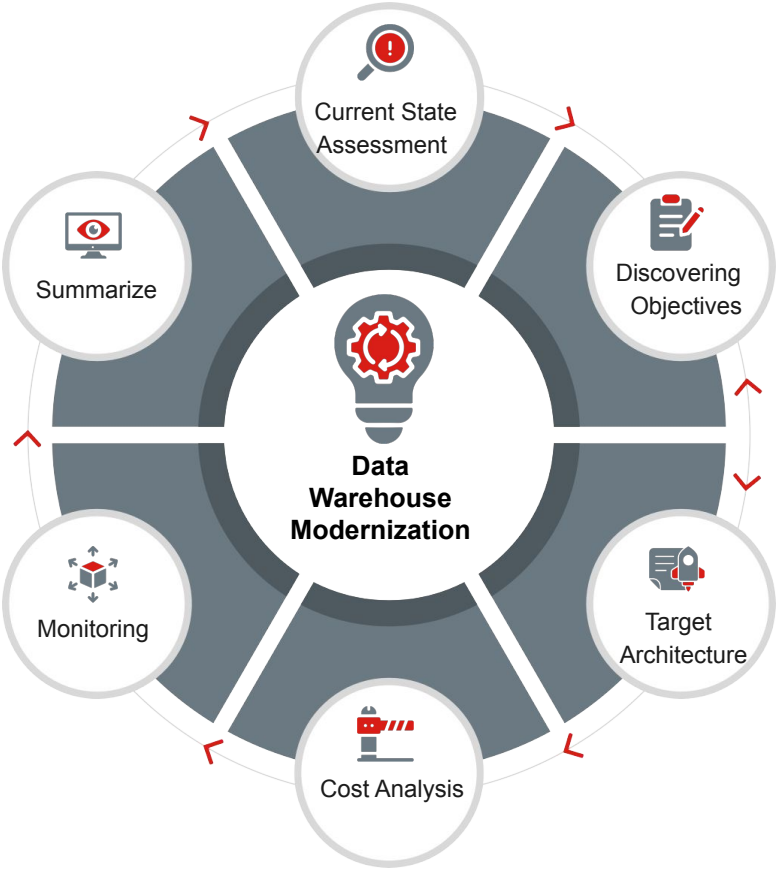
AI/ML & MLOps
Deploy ML models swiftly at scale, and increase return on investment from ML/AI projects.

Business value
Generate insights for business and achieve KPIs using BI, ML models and algorithms.

SIGMOID


Data Warehouse
Modernization

Steps for Data Warehouse Modernization



1. Assessing Current State

- Assessment of the organization's existing data warehouse infrastructure, data models, ETL/ELT processes and analytics capabilities
- Analyzing the query performance, data modelling, indexing strategies and resource utilization

2. Laying down Objectives & Requirements

- Collaborating with key stakeholders for defining the business objectives, priorities, and requirements for warehouse modernization

3. Defining the Target Architecture

- Defining and designing the target data architecture tailored to the customer's specific requirements and workload patterns, following Azure's Well Architected Framework
- Suggesting best practises for data modeling, partitioning, indexing, and workload management

4. Cost Analysis & Optimization

- Performing a detailed cost analysis of the customer's Azure Data Warehouse compute, storage and data transfer costs

5. Continuous Monitoring and Optimization

- Defining the best practises for setting up monitoring and alerting mechanisms for tracking performance metrics and resource utilization
- Providing recommendations for fine-tuning based on changing workload requirements and data patterns

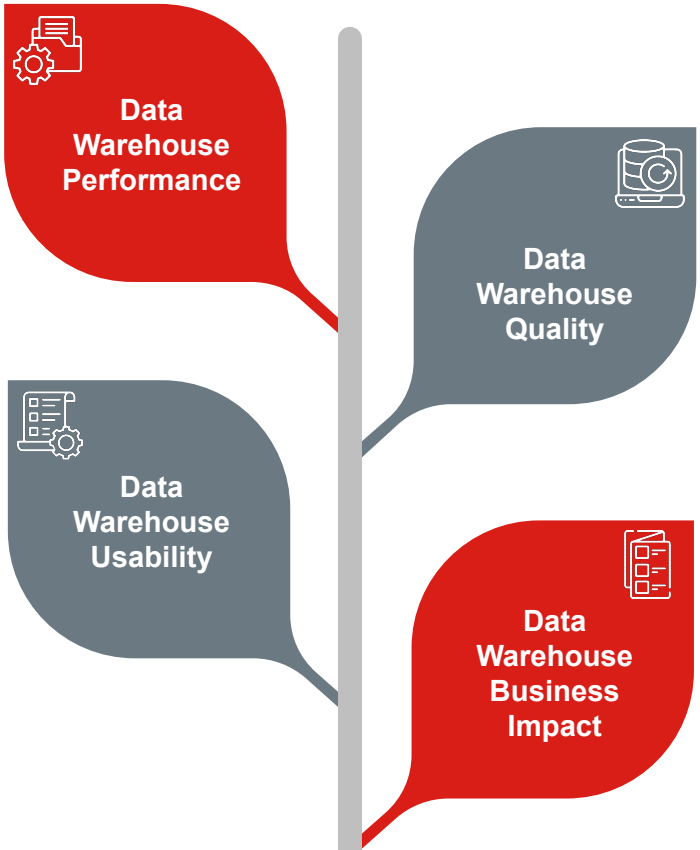
6. Summarize, Train & Knowledge Transfer

- Summarizing the findings of the assessment and presenting actionable recommendations
- Offering recommendations on the best practices for Azure Data Warehouse performance and cost optimization

SIGMOID

KPIs to measure the success of Data Warehouse

- Load time - measures how long it takes to load data from the source systems into the data warehouse
- Query time - measures how long it takes to execute queries and retrieve data from the data warehouse
- Availability - evaluates how often the data warehouse is up and running without any failures or downtime
- Utilization - assesses how much of the data warehouse resources (such as storage, memory, CPU, etc.) are being used and how efficiently



- Accuracy - how well does the data match the source systems and real-world events
- Consistency - does the data conform to established standards and definitions
- Completeness - how much of the data needed to support analysis and reporting is available
- Relevance - does the data align with current and future business needs

- Accessibility refers to how easy it is for users and stakeholders to connect to and access the data warehouse from their preferred tools and platforms
- Understandability measures how easily users and stakeholders can comprehend the structure, content, and meaning of the data warehouse
- Usefulness looks at how well the data warehouse supports analytical and reporting needs
- Satisfaction evaluates how pleased users and stakeholders are with the functionality, performance, quality, and value of the data warehouse

- ROI - how much return on investment have you gained from your data warehouse in terms of increased revenue, reduced costs, improved efficiency
- Alignment - how well does your data warehouse support the strategic objectives and priorities of your organization
- Innovation - how much has your data warehouse enabled or facilitated new insights, opportunities, or solutions for your organization
- Competitive Advantage - how much has your data warehouse given you an edge over competitors in terms of market share, customer loyalty, or brand reputation

SIGMOID

Success Story - Top Fortune 100 retailer achieves near real-time visibility to customer insights with robust data warehousing infrastructure

The client is amongst the world's largest retailers with over 11,000 stores and 245 million customers visiting the stores on a weekly basis.

Case Background

- The existing process of reports generation was extremely manual and intensive, with their business users having to wait for several weeks to finally access them.
- They required an effective data warehousing solution to enable granular analysis on large volumes of data at faster response times, thereby expediting the process of customer insight generation.

Sigmoid's Solution Approach

- Sigmoid's solution enabled ad-hoc analysis on over 250 TB of customer and POS transaction data.
- Enabled slice-and-dice on data to deliver granular customer information such as weekly trends, penetration in terms of customer reach, etc.
- Highly effective data warehousing and management empowered 500+ business users to have full interactive access with a single source of truth (SSOT) to run ad-hoc queries at sub second response times, thereby saving 1000s of end-user hours.
- Completely automated process, removed existing manual labour, inter-team dependencies and delays in accessing data.

< 5 secs
Average Query Response Time

60%
Faster reporting saves 1000s of end-user hrs

250 TB
Data volume processed

Data Sources:

Customer (Experian), POS transactions, Item data, SKU, Store, Calendar information, Weather channel, and others.

SIGMOID

Engagement Models

Sigmoid's Engagement Models

Project Based

- Starts with consulting/scoping (2-3 weeks)
- Delivery Program Management
- Interim review
- Success criteria met and IP handover
- Option to continue with product support
- Fixed bid contract
- 3-5 months duration given complexity of problem

Benefits

- Cost effective
- KPI/SLA/Outcome driven
- Suitable for Fixed scope of work
- Less overheads

Staff Augmentation

- Understanding of skill requirements
- Profile match and rate card
- Onboarding and monthly billing
- Focused training based on client tech stack
- Project Management support
- 10% backup resources unbilled and trained

Benefits

- Scalability
- Flexibility in resourcing
- Ability to change/redefine scope

Hybrid-Flexi Model/Data Labs/CoE

- Mix of project and staff augmentation engagements
- Requirement gathering
- Requirement classification - as project or staff augmentation
- Joint delivery plan
- Secure resources internally from Sigmoid and bill monthly
- Dedicated PM, Engineering Managers
- Dedicated Management Consultant(s)
- Dedicated Team Leads and Product Owners

Benefits

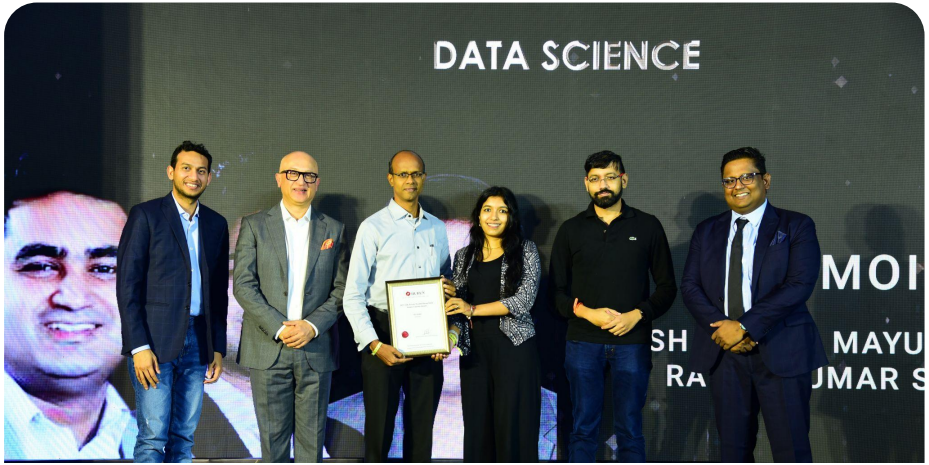
- Cost effectiveness by focus on output
- Ability to change/redefine scope/Change requests
- Risk/Reward linked to KPI/SLA

SIGMOID

Thank you

Email: surabhi.s@sigmoidanalytics.com

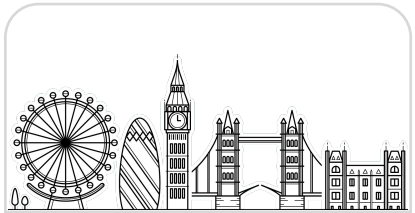
Website: www.sigmoid.com



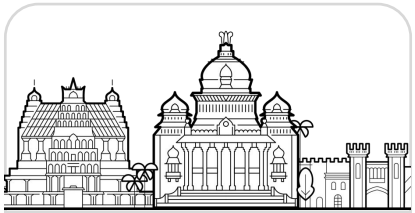
'India Future Unicorn Award'
in Data Science category by Hurun India

Global presence:

USA (NY, SF, Dallas, Chicago)



EU (Amsterdam, London)



India (Bengaluru)

LATAM (Lima)