



 **Microsoft**  
Solutions Partner  
Data & AI  
Azure

# Data Democratization on Azure

[www.sigmoid.com](http://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



**America's  
Fastest-Growing  
Private Companies**



**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

# 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 Partners

-  Microsoft
-  databricks

## Cloud Technologies

-  python
-  PySpark
-  jupyter
-  TensorFlow
-  R
-  NLP
-  MATILLION
-  DATADOG
-  SPARK
-  HADOOP
-  cassandra
-  mongoDB
-  HBASE
-  atlan
-  Alation
-  kubernetes

## Technology Expertise

# 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:

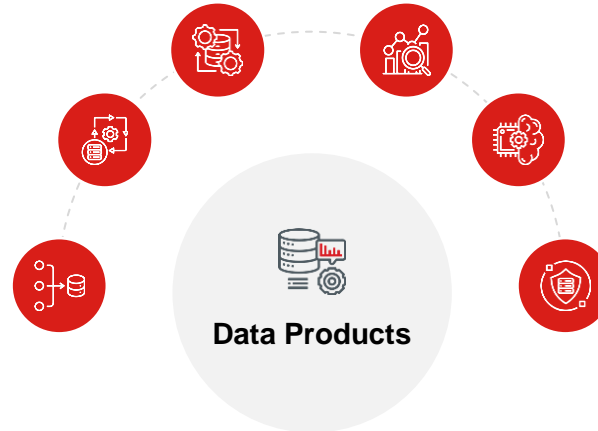
- **Azure Synapse Analytics:** Analytics service which will be used for analyzing large amounts of data using either serverless or provisioned resources.
- **Power BI:** Business intelligence tool to be used for creating interactive visualizations and reports.

## 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.

## 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.



## 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.

## 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.



## Sigmoid's Capabilities - Data Democratization

---

# Sigmoid Capabilities - Experience with data products and data mesh



Sigmoid has extensive experience across 3 large customers in building data products and implementing them as specialized applications or tools that leverage data to provide actionable insights, automate processes, or enable decision-making. We tailor the data products to specific business needs, such as understanding customer behavior or optimizing product performance.



## Data Mesh and Data Products:

Sigmoid's approach is focused on building **data products** and enabling self service through a **data marketplace**. Our goal is to help client build smart data products with integrated and enriched data of high quality that unlocks insights across individual and cross domain datasets and also enables AI/ML through an integrated **ML Sandbox**



Sigmoid's implementation of data products 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 to align the implementation with business goals and ensure success.



Benefits will primarily be in the area of having agility in data product development, scalability, improved data quality, and alignment with business goals.

Data products and data mesh are powerful concepts that would enable the client to leverage data for strategic advantage. Implementing these strategies would require careful planning, alignment with business objectives, and a focus on data quality and governance which Sigmoid promises to.

# Bringing Data Democratization through the implementation of Data Mesh

The decentralized data architecture approach of data mesh facilitates the creation of curated, easily accessible domain oriented as well as cross domain data products that ensure interoperability, data discovery and self serve analytics

## Advantages realized by our customers:

**Agility - Faster development and iteration of data products.**

**Quality-Higher data quality due to domain expertise and ownership.**

**Efficiency - Reduction of bottlenecks associated with centralized data management**

**Innovation: Has brought experimentation and innovation within teams**

## Foundational Data Products

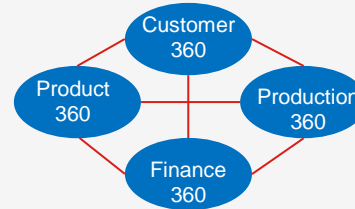
### Curated Data

Data assets that are curated by **domains** and form **the source of truth** on top of which additional data assets are built

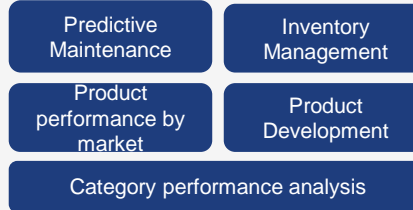


### Integrated Data

Data assets that are **cross functional** and **cross domain**, and serve multiple business use-cases



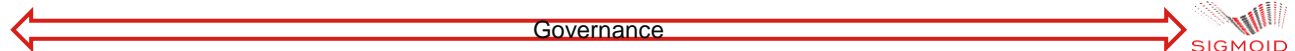
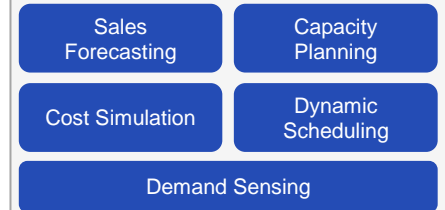
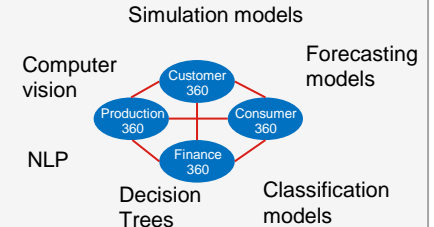
### Illustrative Use Cases



## Smart Data Products

### Enriched Data

**Data Science/AI algorithms** are applied to data in order to enrich & generate insights









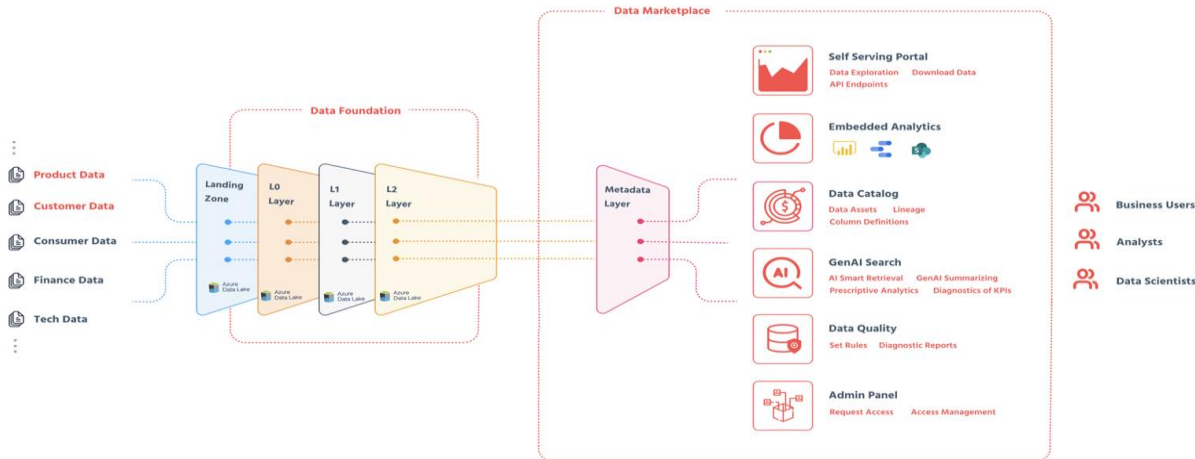
# Facilitating Data Discovery and Efficient Usage through Data Marketplace

## What is a Data Marketplace

An internal data marketplace is a centralized platform within an organization that facilitates the discovery, sharing, and governance of data assets across various departments and teams. It operates much like a commercial data marketplace but is tailored for internal use, enabling employees to access and leverage data efficiently to drive business insights and decision-making

## Typical Problems Solved by a Data Marketplace

-  **Siloed Data**
-  **Delays in accessing needed information**
-  **Inefficient usage of available data**
-  **Data inconsistency**

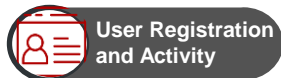


## Factors taken into consideration:

**A precursor to building a data marketplace for an org is to determine the strategic goals and organizational maturity.**

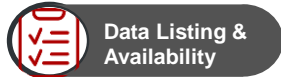
- Stage of data maturity: Ideally the processes should be standardized and documented
- Identified Data Challenges and Needs
- Strategic Alignment of the Organization's Goals
- Should have necessary technological infrastructure and integration capabilities
- Cultural and organizational readiness for successful adoption

# KPIs to Measure Data Marketplace Success



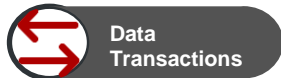
**User Registration and Activity**

**Number of Registered Users:** Measure the total number of users who have registered on the data marketplace.  
**Active Users:** Track the number of users actively engaging with the platform over a specific time-period.



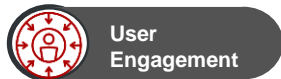
**Data Listing & Availability**

**Number of Data Listings:** Monitor the quantity and diversity of data listings available on the marketplace.  
**Data Source Diversity:** Assess the variety of data sources contributing to the marketplace.



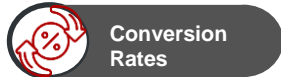
**Data Transactions**

**Number of Transactions:** Count the total number of data transactions, indicating the level of engagement.  
**Transaction Volume:** Measure the amount of data being transacted, either in terms of volume or value.



**User Engagement**

**Time Spent on Platform:** Monitor the average time users spend on the data marketplace.  
**Frequency of Use:** Track how often users return to the platform.  
**Self-Service Usage:** No. of new self-service reports created



**Conversion Rates**

**Conversion from Visitor to User:** Measure the percentage of visitors who become registered users.  
**Conversion from User to Customer:** Track the percentage of users who make a transaction.



**Customer Satisfaction**

**User Ratings and Reviews:** Collect feedback for data quality and overall satisfaction.  
**Customer Support Metrics:** Monitor response times, issue resolution rates, and overall customer support effectiveness.



**Data Quality Metrics**

**Accuracy and Reliability:** Assess the quality and reliability of the data available on the marketplace.  
**Data Update Frequency:** Measure how frequently data listings are updated.



**Marketplace Reach**

**Geographic Reach:** Evaluate the geographic diversity of users and data sources.



**Platform Health**

**Platform Uptime:** Ensure the platform is consistently available and operational.  
**Technical Performance:** Monitor load times, response times, and overall technical performance.

**Realized Results with a CHC major**

**60%**  
faster implementation of new use-cases

**65%**  
wider scaling of business results

**20%**  
growth in sales lift

**1.5 MM**  
reduction in tech debt

# Requirements for Data Marketplace

## Business Requirements



Create a scalable and performant platform for creating, consuming and sharing data products/data assets from all data domains of Product, Customer, Consumer, Supplier, People and Finance. This platform should be the single hosting point for all Client data products including cross domain and integrated data products.



Provide self service functionality for data exploration and data set creation from L1, L1+ and L2 layers, with GenAI based and responsive search options to find and consume data. Including ability to filter and sort data by attributes like data type, owner and creation date.



Provide lineage, table and field level Business/Technical description, metadata and data quality metrics.



Support workflow for data governance and requesting access to data.

## Business Requirements



Support Findability, Accessibility, Interoperability, Reusability (FAIR) principles.



Provide a intuitive user journey for different user types and integration with Analytics/BI tools. Should also have in built help, documentation and training.

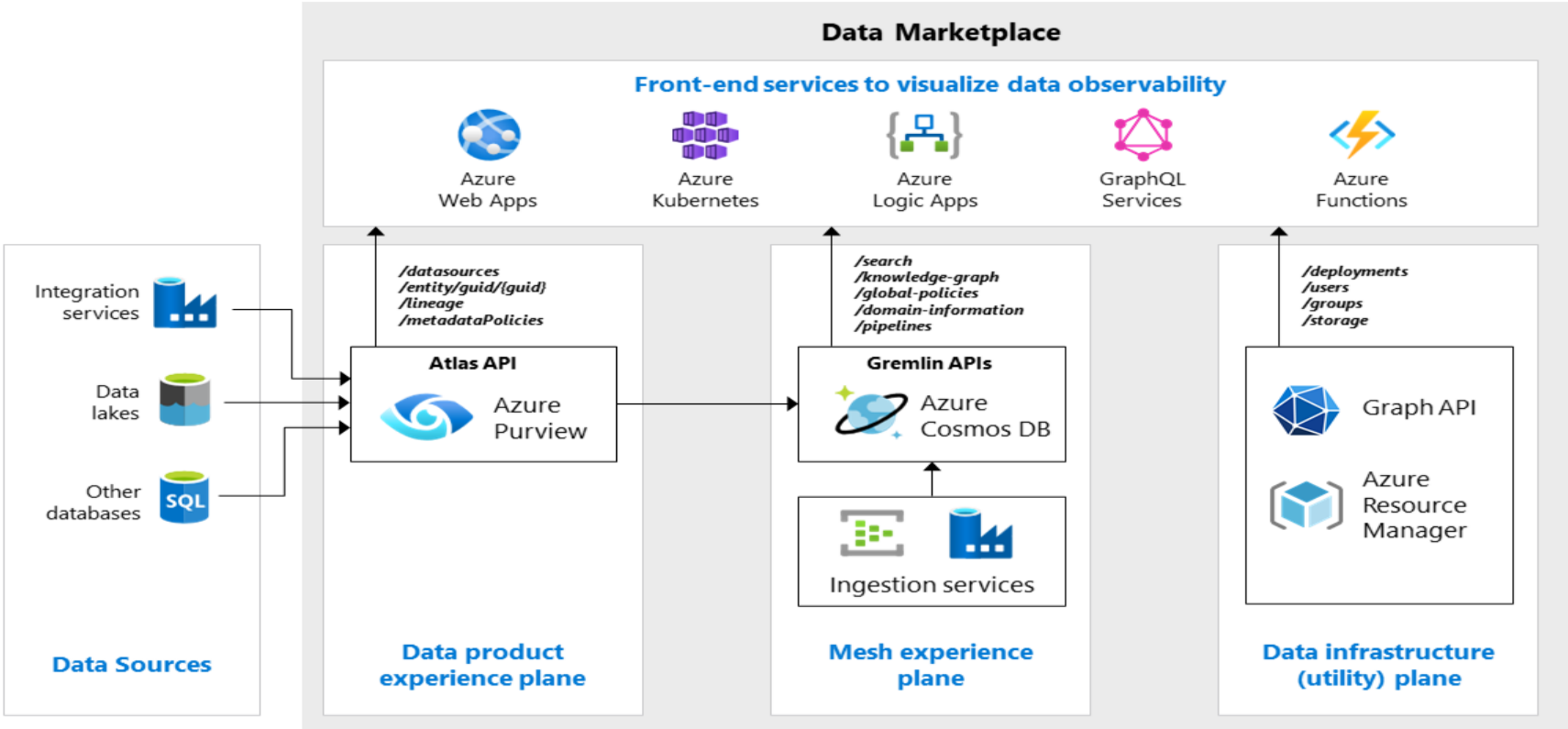


Should support all defined identity, access management policies and also provide logging and audit capabilities



Provide additional capabilities like - Versioning, API based Integration and Rich visualizations.

# Recommended Enterprise Architecture of a Data Marketplace on Azure





## Success Story

---

# Sigmoid's Work in Data Marketplace

We have built data marketplaces to facilitate the monetization and acquisition of data and data catalogs to enhance data discovery, governance, and collaboration within an organization. As an example for a customer we created **supply chain data products** and canonical data models

## Problem:

The customer, a F500 firm was facing inventory management issues due to lack of consistent, quality data and information transfer between multiple user systems. Consolidation of data across disparate systems was also a challenge

## Sigmoid's Solution:

- Sigmoid modernized the data architecture by creating a data mesh driven architecture, enabling the creation of data products for multiple business domains with greater ownership to the users
- The custom-built data connectors were developed using a Low-Code No-Code methodology to streamline the integration of multiple existing data sources with streaming data sources

The Data Marketplace had the following key features:

1. **Data Discovery:** Allowing users to search and explore various datasets available drill down and analysis.
2. **Data Quality Assurance:** Ensuring that the data meets certain quality standards as defined in the DQ tool Collibra. These rules included business rules apart from technical rules.
3. **API Integrations:** Providing API interfaces for downstream enterprise applications to consume the modeled data.
4. **Query Interface:** Providing access to Azure delta live tables for querying and data downloads.

## Benefits:

40%  
increase in on-time  
product delivery

60%  
higher data usability  
across domains

## Other Use Cases enabled for users:

1. **Data Discovery:** Helped data analysts, business users and scientists find the data they need for their work.
2. **Data Governance:** Supports data stewardship and compliance by maintaining a clear record of data assets and their usage.
3. **Knowledge Sharing:** Facilitates collaboration and knowledge sharing among data professionals within an organization.

# Engagement Model

---



# 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/define 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/define scope/Change requests
- Risk/Reward linked to KPI/SLA



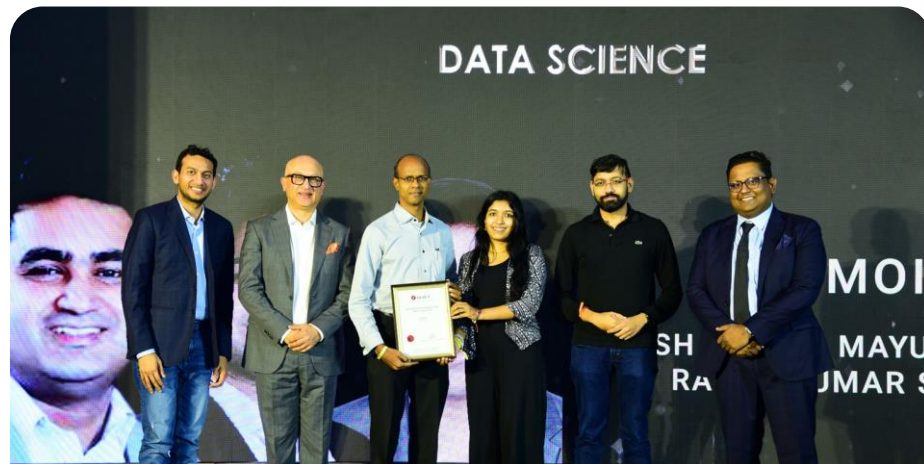
# Thank you



Email: [surabhi.s@sigmoidanalytics.com](mailto:surabhi.s@sigmoidanalytics.com)



Website: [www.sigmoid.com](http://www.sigmoid.com)



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

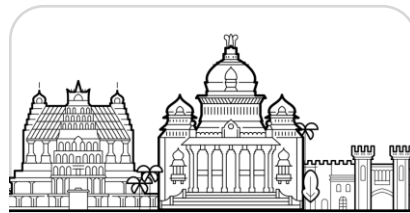
## Global presence:



**USA** (NY, SF, Dallas, Chicago)



**EU** (Amsterdam, London)



**India**  
(Bengaluru)



**LATAM** (Lima)