

# Beam Data

Consulting Service Introduction

# About Beamdata

- **Beam Data** is a Toronto based consulting company that provides full lifecycle Data and AI consulting services to clients in various industries.
- Whether you are an early-stage startup that needs help with data strategies or a mid-sized company that needs project resources on building end-to-end machine learning solutions, Beam Data is here to help tackle your most challenging problems.
- We also provide the customized corporate training services on Python, Data Science, Machine Learning, Cloud Computing, Big Data and workshops for companies of all sizes. Our past clients include major banks, retail, telecommunications, tech startups, and human resources.

# Our Clients

e-Commerce

FinTech

Telecom

Consumer Electronics

Supply Chain

Tech Startup

Media & Publishing

Information Technology



SAMSUNG



COMMUNITECH



THE WORLD'S GREATEST COLLECTOR CAR AUCTIONS™

TOROMONT



ENKIDOO<sup>ai</sup>



Instant Funding  
Your Loan Approved On Time



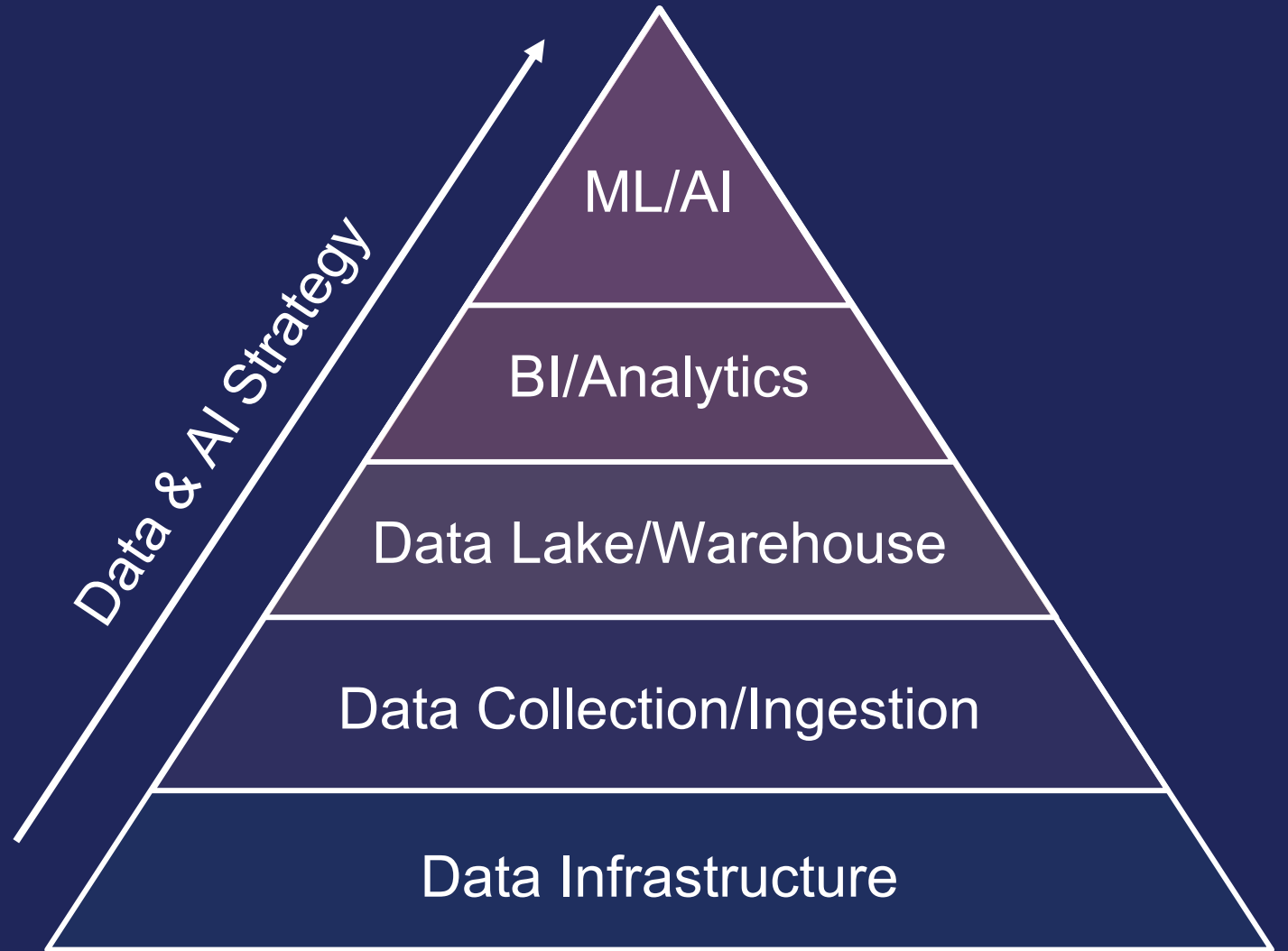
FURNISHR.



sponsorcircle

EX-A-R  
STUDIOS

# Our Services

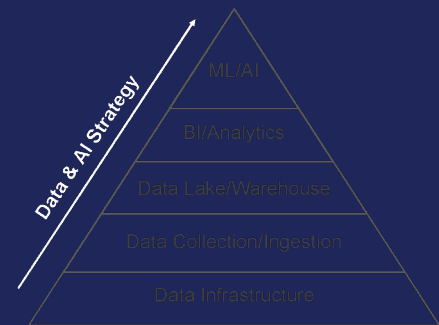


# Data Strategies

We provide data strategy consulting to clients from early-stage startups to enterprises going through digital transformations.

## Services include

- Building the short/long-term data and AI roadmap
  - Help build a data roadmap to meet company's growth needs
  - Help shape core data strategies that fit into the product roadmap
- Data capability assessment
  - Help assess the current data/AI capabilities and find areas of improvements
  - Pinpoint the areas that data can bring real business value
- Executive team training and data team hiring strategy
  - Provide executive-level AI/data literacy training
  - Create the most fit hire vs build strategy

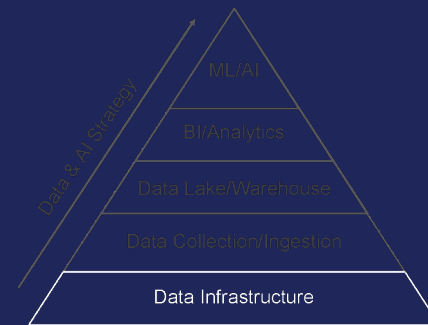


# Data Infrastructure

We help our clients evaluate, plan, and build data infrastructure that is secure, robust, scalable, economic, and portal.

## Services include

- Data Infrastructure Assessment
  - Help assess current data infrastructure and suggest best actions for modernization
- Cloud Data Migration Planning
  - Data infrastructure tooling research and benchmarking
  - Creation of the cloud migration plan
  - Security and compliance check
- Data Infrastructure Implementation
  - Cloud infrastructure set up (DCs, Networking, IAM, etc.)
  - Data migration to the cloud (Azure, AWS, and GCP)
  - Best practices in data security and protection
  - Infrastructure as Code, DataOps, and Automation

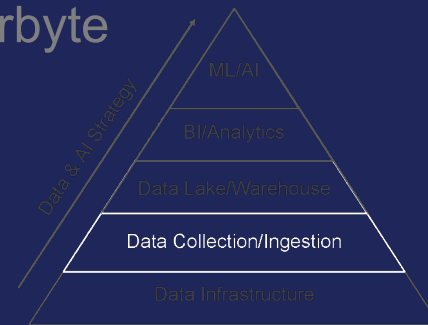


# Data Collection & Ingestion

We help our clients build first- and third-party data collection strategies and implement data ingestion pipelines.

## Services include

- Data collection strategy
  - Help assess first- and third-party data collection needs
  - Help build the CDP/CDI strategies for MarTech and the Cookieless world
  - Data source research and feasibility analysis
- Data collection implementation
  - Help collect data through web data APIs, web scraping, and other means
  - Implement Tag Management and CDI data collection pipelines
  - Create custom data connectors for data ingestion
  - Help deploy lowcode connectors such as Fivetran or OS tools such as Airbyte
- Data annotation (Machine Learning)
  - Help with manual data labelling on a smaller scale
  - Work with 3<sup>rd</sup>-party labelling platforms to label/annotate data at scale

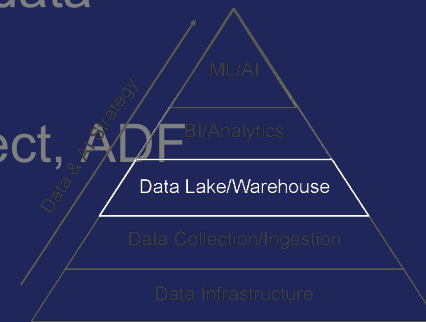


# Data Lake/Warehouse

We specialize in modernizing enterprise data warehouse and building data lake/house for our clients

## Services include

- Data Integration Service
  - Build ETL/ELT pipelines for data warehouse/lake (Spark, Snowflake, dbt, etc.)
  - Big data processing using Apache Hadoop and Spark (Databricks, AWS EMR)
  - Data quality governance
- Data Warehouse & Lakehouse
  - Build Cloud-based Data Warehouse using the Modern Data Stack
  - Architect and implement the Data Lakehouse solutions for streaming data
- Data Pipeline Orchestration
  - Orchestrate and automate complex data pipelines using Airflow, Prefect, **ADF**



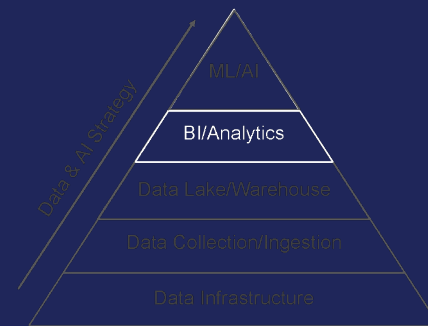


# BI & Advanced Analytics

Our data analytics experience spans many industries and use cases

## Services include

- BI Dashboards
  - Architecting data models for BI Analytics
  - PowerBI/Looker/Tableau dashboards for business teams
- Business forecasting
  - Time series forecasting for various business use cases
- Customer analytics and predictive modeling
  - Customer acquisition models
  - Customer segmentation
  - Customer lifetime value prediction
  - Predictive churn modeling for customer retention
  - Retail product assortment
  - etc.

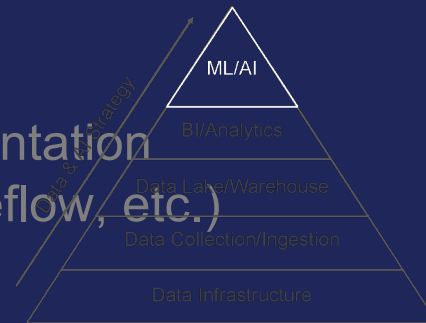


# Machine Learning & AI

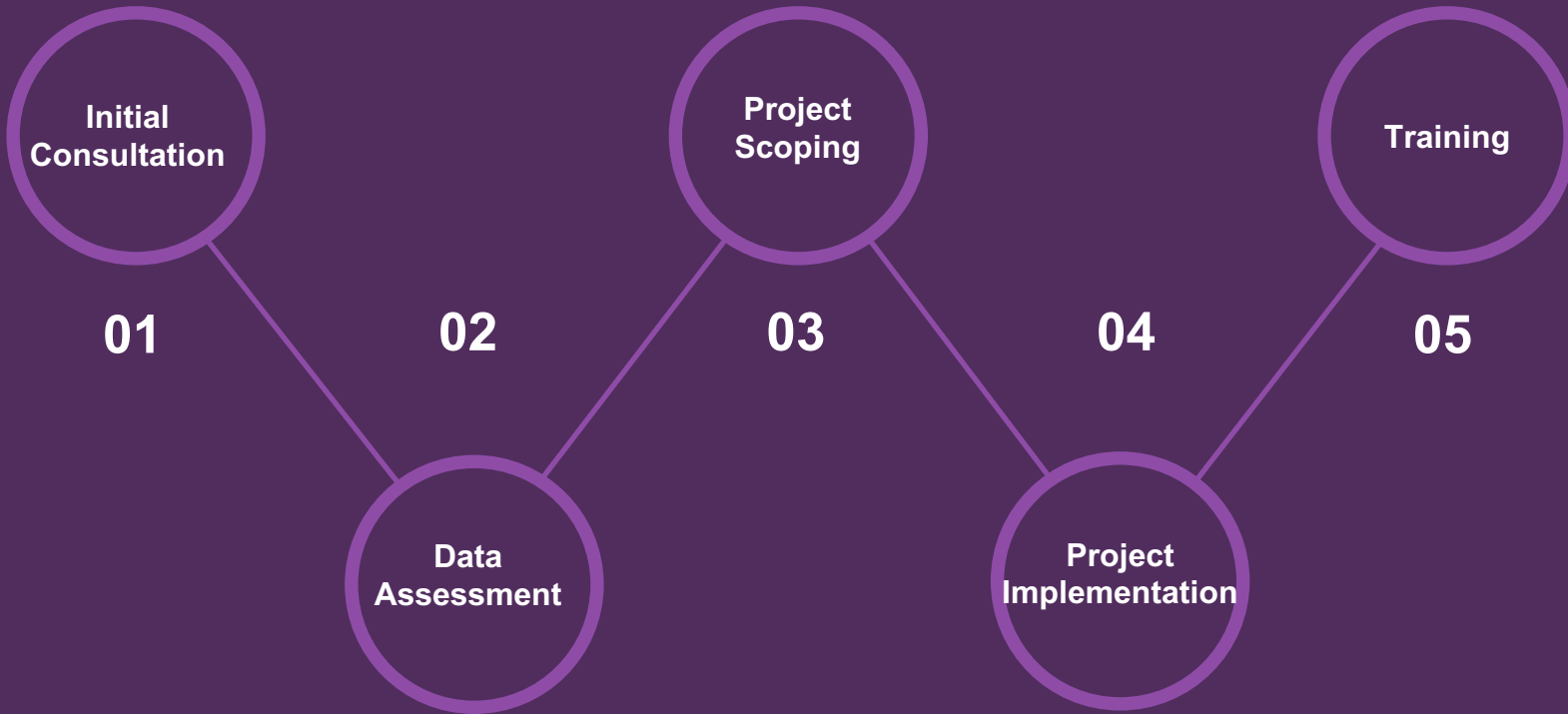
We provide full lifecycle ML services from experiment to production

## Services include

- Machine Learning & AI
  - Help clients build high-level machine learning strategies
  - Support clients with AI/ML Model Tuning & Optimization
  - Covering the following use cases
    - General Classification/Regression problems
    - Natural Language Processing (Sentiment Analysis, Topic Modeling, NER, Text Summarization, etc.)
    - Computer Vision (Image Classification, Object Detection, Semantic Segmentation)
    - Recommender Systems (News, Product, Social)
    - Fraud/Anomaly detection
- MLOps
  - Help clients build MLOps strategies
  - End-to-end model training, serving, deployment, and monitoring implementation
  - Model deployment with different services (SageMaker, MLflow, and Kubeflow, etc.)



# Working with Beam



# Working with Beam



**01**

## Initial Consultation

Try to pinpoint the business/data challenges

## Data/Model Assessment

Get to know client's current data stack and data models

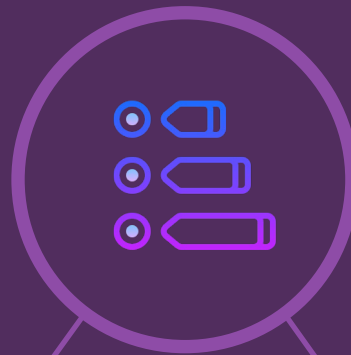
**02**



**03**

## Project Planning

Create a detailed and solid plan for implementation



**04**

## Project Implementation

Work on execution to achieve the goals according to the plan



**05**

## Final Delivery

Deploy services in production, document everything, work on transfer and provide end-user training



# Working with Beam



## Initial Consultation

Before we start a project, we usually arrange a few consultation sessions to understand our client's expectations. In the initial consultation sessions, we not only want to know the business pain points, but also want to understand our client's expectations.

# Working with Beam

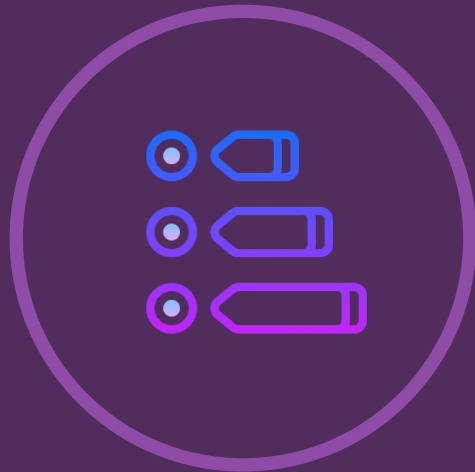


## Data Assessment

Knowing our client's existing data capability is very important. Beam Data will meet with the client's IT and Data teams to do an overall assessment. In this survey, we are hoping to understand the following:

- Data/IT team structure
- Current data infrastructure
- Database models
- How existing data pipelines are orchestrated
- Existing ML models and high-level intro

# Working with Beam



## Project Planning

Based on the data assessment and client's business requirements, we will break down the bigger goals into smaller work streams.

- Milestones will be set for each work stream
- Resources will be planned according to the scope of the work and timeline

We make sure that the deliverables are clearly defined so both parties can agree on when a milestone is considered as *DONE*.

# Working with Beam



## Project Implementation

Beam Data's project implementation team will be spending majority of the time on actual implementation and delivery. We use JIRA or client's Kanban tools for project task issue tracking and management.

- Project members on the client's side will be kept in the loop on a regular basis and we will meet for weekly/bi-weekly project review and planning
- We iterate fast and client will be able to see the results and help with UAT so suggestions can be incorporated into the next cycle



# Working with Beam



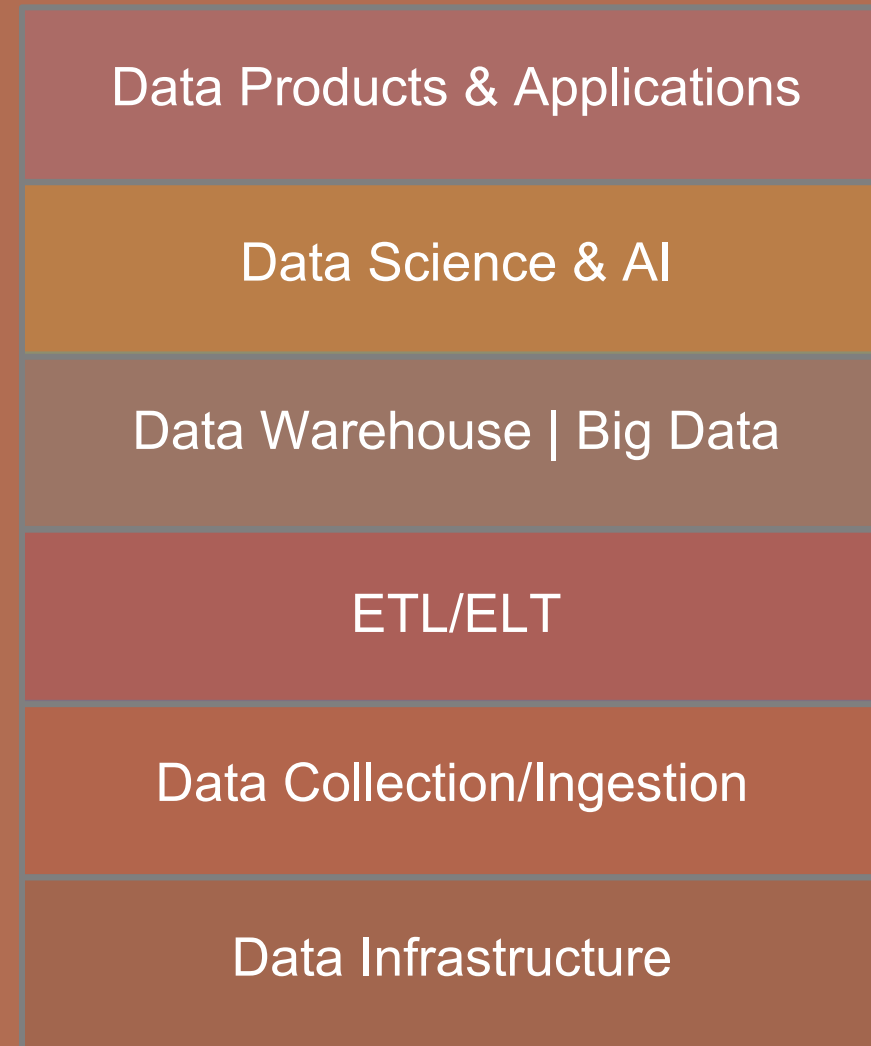
## Final Delivery

Before a project is wrapped up, we will spend time documenting the entire project and the most important processes, diagrams, workflows, models, and algorithms.

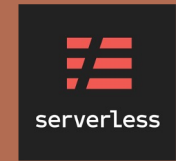
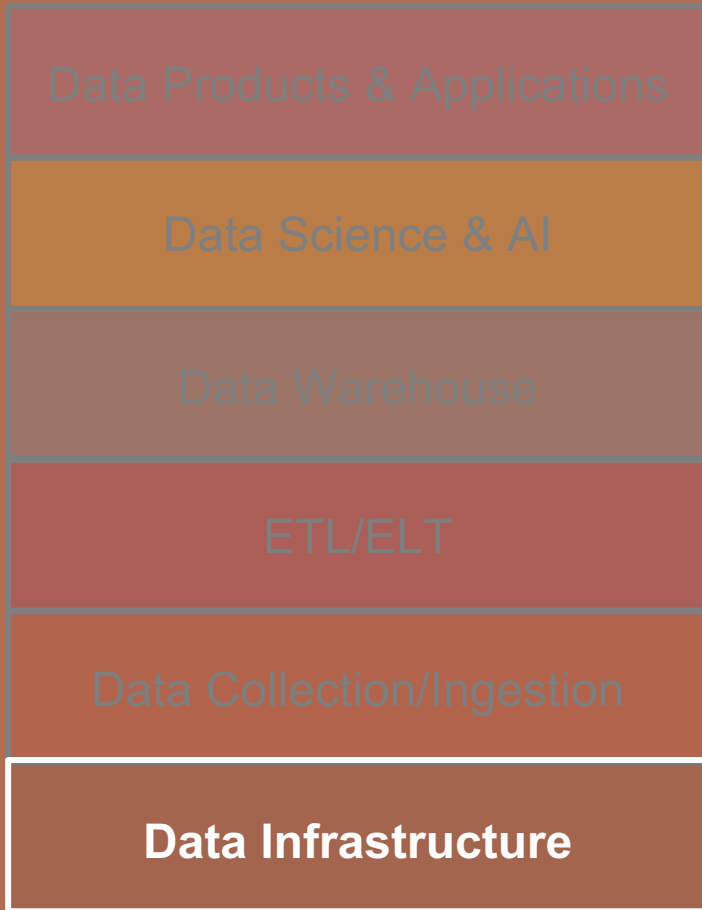
A final presentation will be booked and client's Data/Product/ML teams will join to get an update on the final data product.

We will then arrange staff training to make sure that the two project teams have a smooth transition.

# Modern Tech Stack We Use

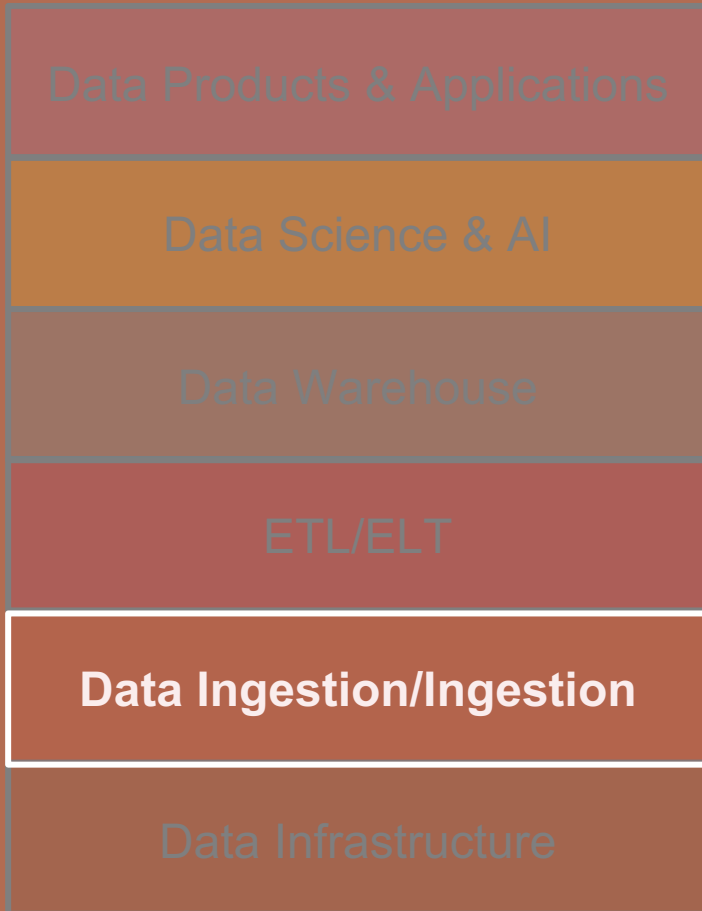


# Our Tech Stack



- Our team is familiar with all 3 cloud platforms: AWS, Azure, and GCP
- We have helped our clients provision and deploy Infrastructure as Code (IaC) via CloudFormation, CDK, and Terraform
- We have supported our clients with Kubernetes and Kubeflow platform implementation. We also deliver Kubernetes training to our clients in the Financial Industry
- If Serverless is an ideal use case, we have experience developing and deploying serverless applications via Lambda and Azure Functions

# Our Tech Stack

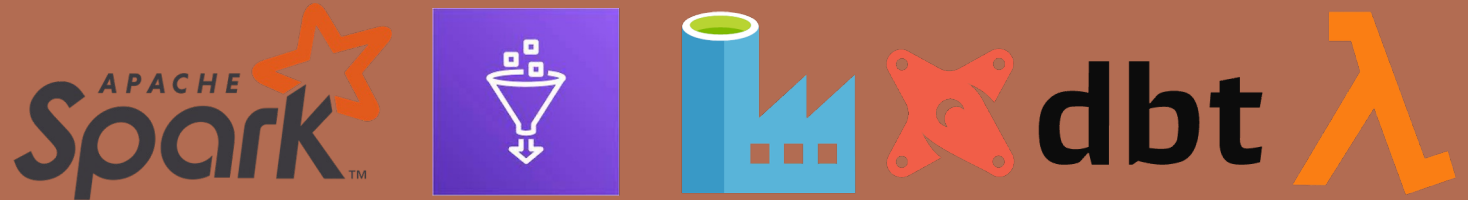
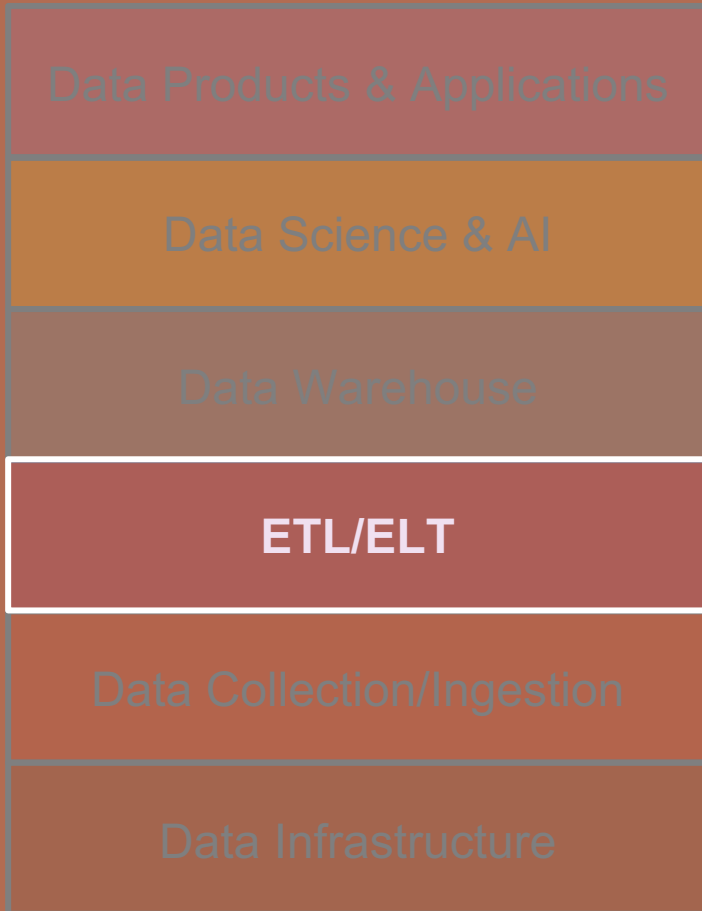


nifi



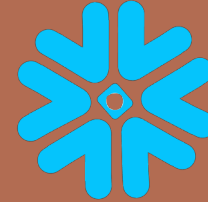
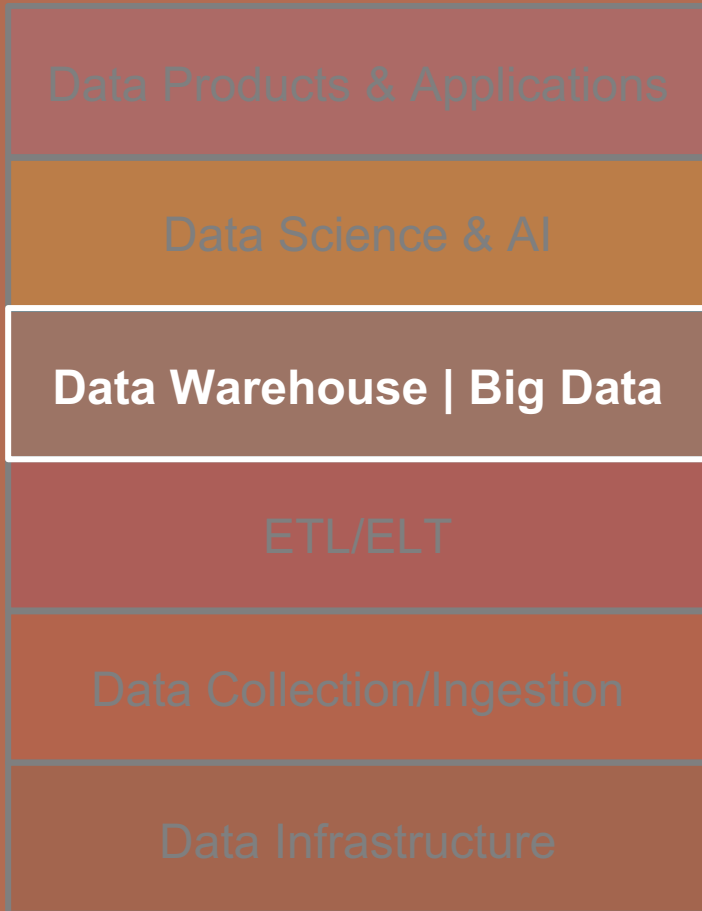
- Data collection and ingestion is a key component of any data project.
  - We have helped many clients run web scraping using Python toolkits, data collection from Data APIs such as Twitter, Instagram, Google Analytics, etc. for various product analytics and market research use cases using the Python selenium and scrapy packages.
  - We also have experience ingesting streaming data using Kafka/Kinesis using the Delta Lake or Apache Hudi/Iceberg formats
- We also have expertise in setting up low-code data ingestion tools such as Airbyte and FiveTran to help clients deliver data to lake and warehouse rapidly

# Our Tech Stack



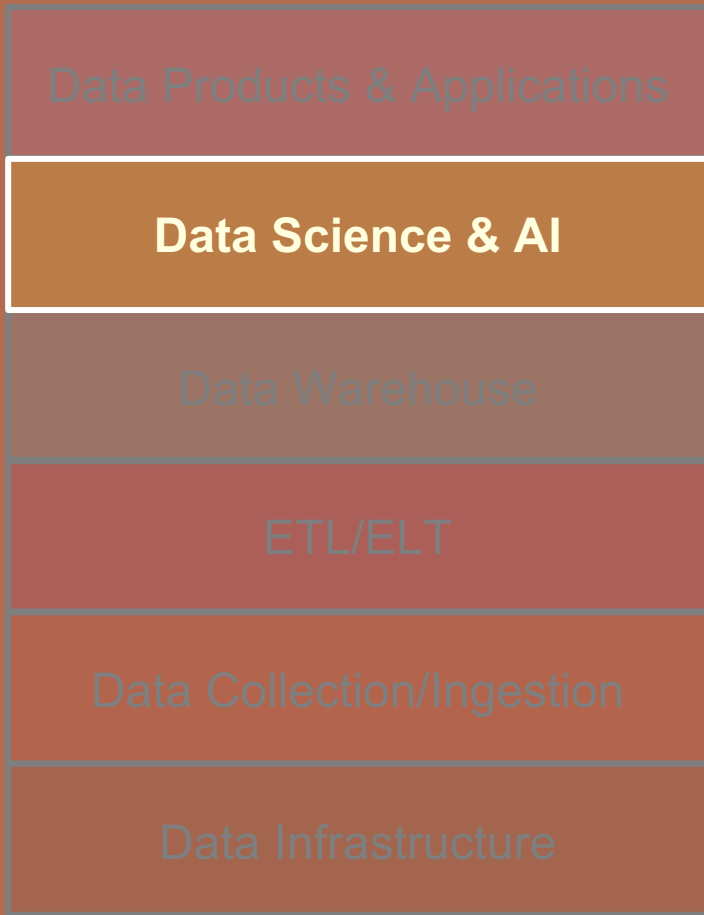
- When it comes to data integration, we have experience in the traditional ETL approach, the modern ELT approach, as well as the serverless framework. We help our clients adopt the most suitable approach and tools to meet their needs
  - Developing and executing Spark ETL jobs on dedicated Databricks clusters, Amazon Glue jobs, and Azure Synapse Spark Pools
  - Building low-code transformations in Azure Data Factory
  - Running dbt workflows for ELT in data warehouses
  - Building and deploying serverless ETL pipelines via Azure Functions or AWS Lambda
- Beam Data can help our clients choose the stack that not only fits their needs, but also makes economic and scalability sense as the service scale

# Our Tech Stack



- We have expertise in strategizing and implementing modern data warehouse solutions on major cloud platforms such as AWS, GCP, Azure, as well as cloud agnostic platforms like Snowflake
- We specialize in data modeling and big data analytics
  - Dimensional modeling
  - Slowing changing dimensions
  - Building wide tables and data marts
  - Fast queries with TrinoDB, Athena, and Dremio
  - Warehouse performance tuning for Snowflake, Synapse, and Redshift

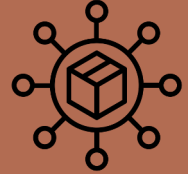
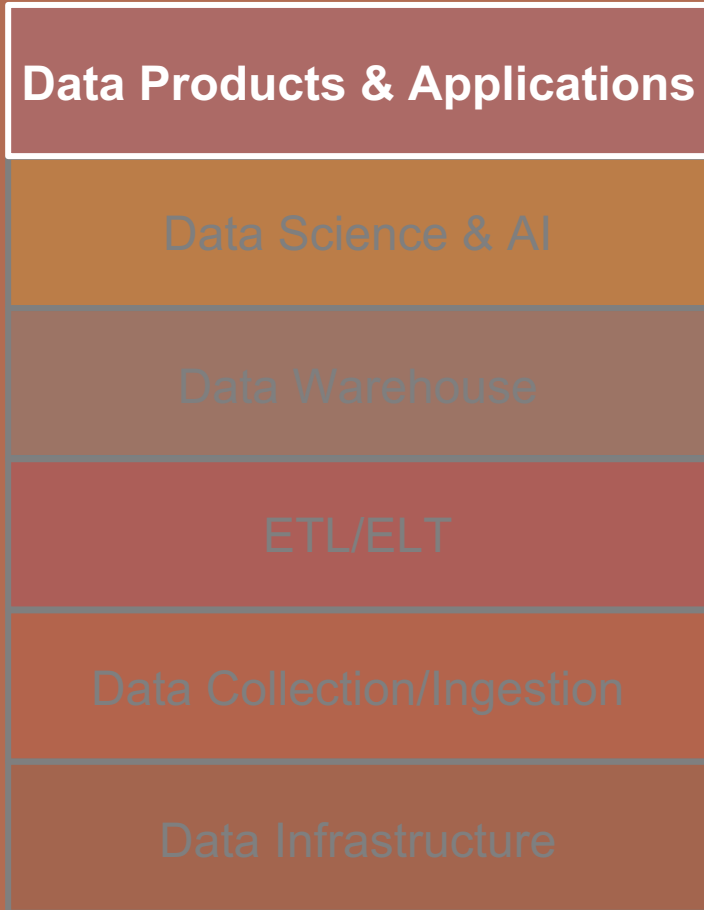
# Our Tech Stack



Data Engineering	Model Engineering	Software Engineering	DevOps
<p><b>Data Labeling</b>      <b>Feature Store</b></p>	<p><b>Experiments</b>      <b>Registry</b></p>	<p><b>Development</b></p>	<p><b>Version Control</b>      <b>Workflow Automation</b></p>
<p><b>Data Query</b></p>	<p><b>Model Tunings</b></p>	<p><b>Model Testing</b></p>	<p><b>CI/CD</b>      <b>Monitoring</b></p>
<p><b>Data Governance/Quality</b></p>	<p><b>Model Training/Evaluation</b></p>	<p><b>Model Serving</b></p>	<p><b>ML Infrastructure</b></p>
<p><b>Data Lake/Warehouse</b></p>			

- Beam Data has worked with many ML tools in the past. We are flexible with the ML tools required to get the work done. We can help our clients anticipate the future tooling needs and choose the right tools
- We have helped our clients deliver many AI projects by leveraging tools like Tensorflow, PyTorch, ONNX, TFServing, to managed services like SageMaker, Kubeflow, and Azure ML

# Our Tech Stack



- Building a scalable and robust data product is much more than just training the best machine learning models and data models.
- Beamdata has experiences in building Data APIs, deploying and serving models in production
- We also specialize in insight discovery and building dashboard products using the most popular data visualization tools such as PowerBI and Tableau



# Book a Consultation

Website: <http://beamdata.ai/>

Tel: +1(647) 459 8172

Email: [info@beamdata.ai](mailto:info@beamdata.ai)

Address: 433 Yonge St 2nd Floor, Toronto, ON M5B 1T3