Articulating Enterprise Data Strategy

E. Doro PhD

Director - Data Science Walmart USA

Outline

- Enterprise Data Strategy + Analytics
- Setting Business Analytics Agenda
- AI Strategy StoryBoarding: Enablers & Critical Success Factors

Connecting Data Strategy to Analytics

Enterprise Data Strategy + Analytics

- **Enterprise data strategy** derives from the **org's business** strategy.

- You get a sense of your **data strategy** as you think through questions like:
  - Why is the **company** in business?
  - What is our **core strength**?
  - Which **customers** should we continue to serve or start serving?
  - Which **products**/**services** should we stop offering, continue to offer or start offering?
  - Why have we decided on these strategic **directions**?

Enterprise Data Strategy + Analytics

- **Data strategy** must be defined **within context** of the larger business strategy.

- A **road-map** for an organization's potential to harness data-dependent capabilities.

  - Umbrella for all **domain-specific strategies**, e.g. master data management, business intelligence

- Good enterprise data strategy is:

  - Practical, Relevant, Evolutionary and Integrated

Enterprise Data Strategy + Analytics

- Formulating **enterprise data strategy** correctly means providing answers regarding:
  - Data governance
  - Datasets
  - Models and AI products
  - Process
  - Infrastructure
  - Talent

Enterprise Data Strategy + Analytics

- **Data Governance**
  - What is the ideal **management structure** for data governance given the org's business strategy?
  - Who owns **formulation**, **execution** and **enforcement** of data policy and guidelines?
  - What part of the **org chart** should the data science group be situated?
  - How does the data team **interface** with other teams across the org?
  - How do we account for **investments** in data systems?
  - What is the process for resolving **data conflicts**?

Enterprise Data Strategy + Analytics

- **Datasets**
  - What datasets do we need to execute the company's business strategy?
    - i.e. what types of data should we collect internally and or externally to answer the business questions and proposals outlined in the enterprise business strategy?

Enterprise Data Strategy + Analytics

- **Models & AI Products**
  - For the datasets collected to execute our business strategy, what **machine learning** models or **business intelligence** products are required?
  - What **metrics** do we need and how do we track and visualize?
  - How do we **integrate** the machine learning products built into our day to day business and tech operations?
  - What types of **experiments** should we run (observational, randomized)?
  - Should we combine models with experiments?

Enterprise Data Strategy + Analytics

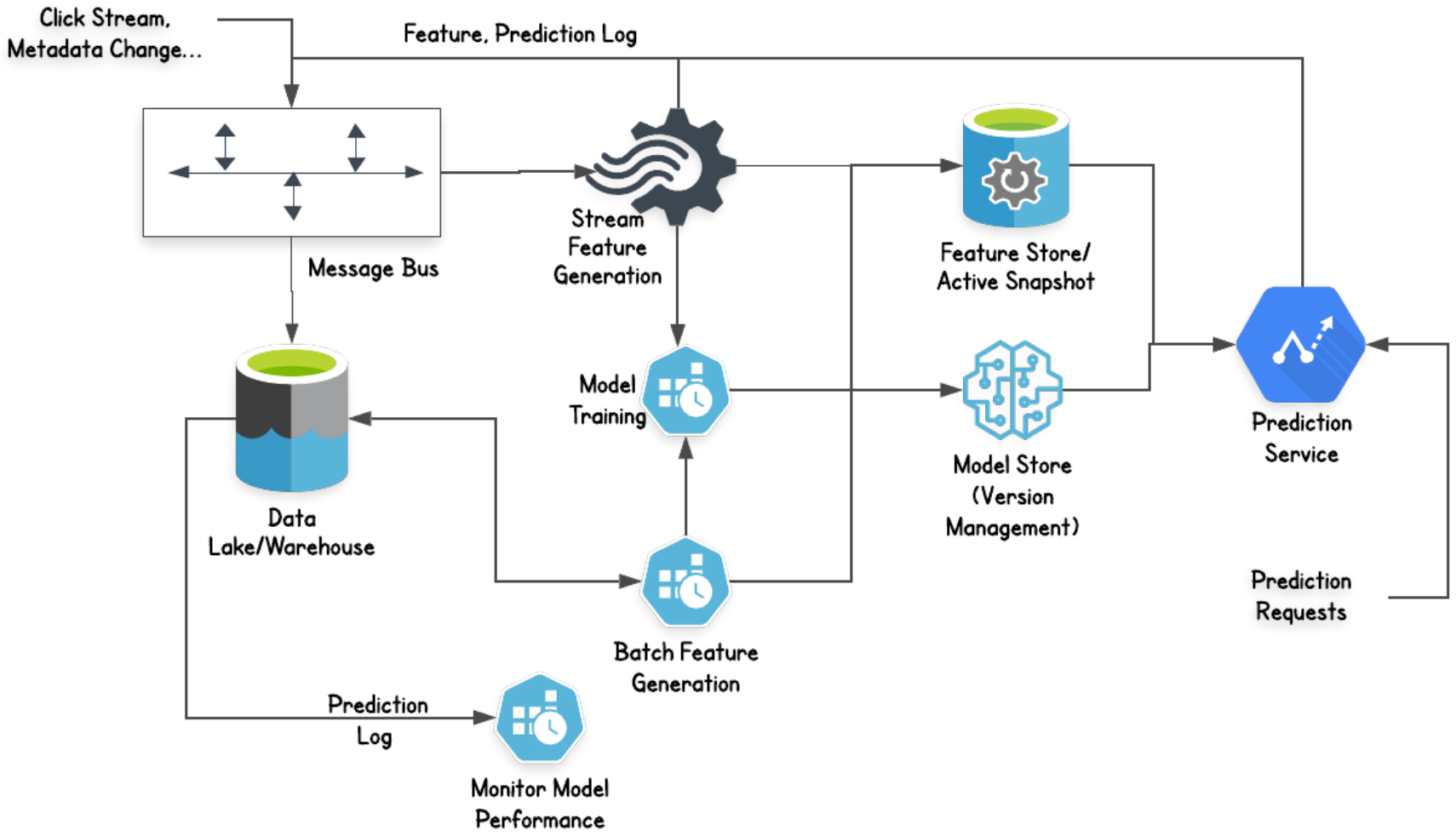
- **Process**
  - How do we ensure data **integrity**?
  - How do we guarantee data **security**?
  - How do we guarantee a **unique global source of truth** for our data across the entire org?
  - How do we provide **access** to data and data-products while preventing misuse or abuse?
  - Would model training, validation and productionizing happen in batch or stream **modes**?
  - How do we implement **model governance**?
  - How do we formulate, deploy, analyze **experiments**?

Enterprise Data Strategy + Analytics

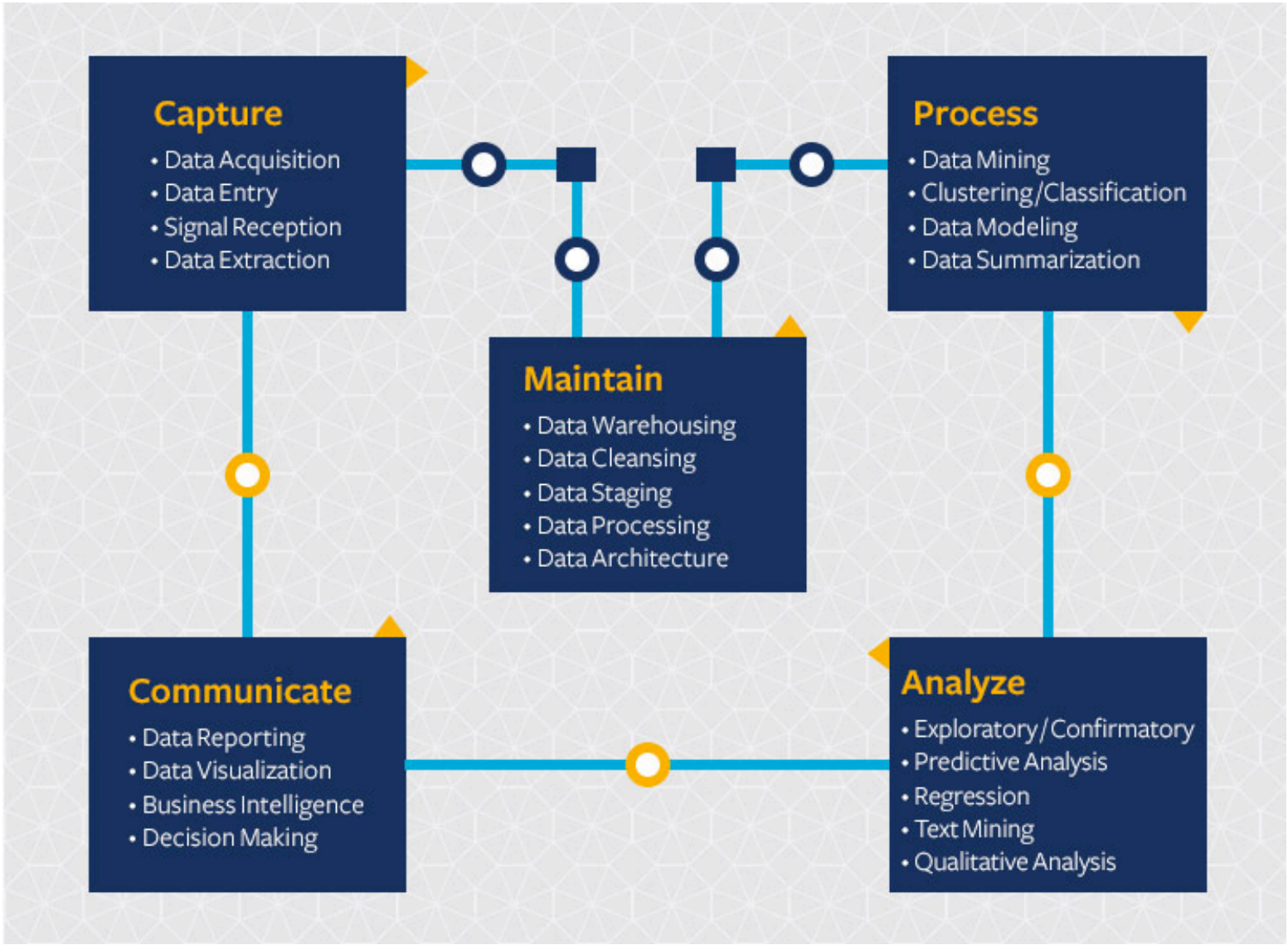
- **Infrastructure**
  - Given our process requirements, what infrastructural support do we need?
  - Should infra be built **in-house** or sourced **externally**?
  - Should we own hardware or employ cloud-based systems?
  - How do we maintain and update these systems?
  - What happens when systems **fail**?
  - Does infra support **scaling** and **increasing complexity** of data and data products?
  - How **adaptable** is our platform?

Enterprise Data Strategy + Analytics

- **Conceptual Enterprise DS Infrastructure**



Enterprise Data Strategy + Analytics



ETL Life Cycle

Enterprise Data Strategy + Analytics

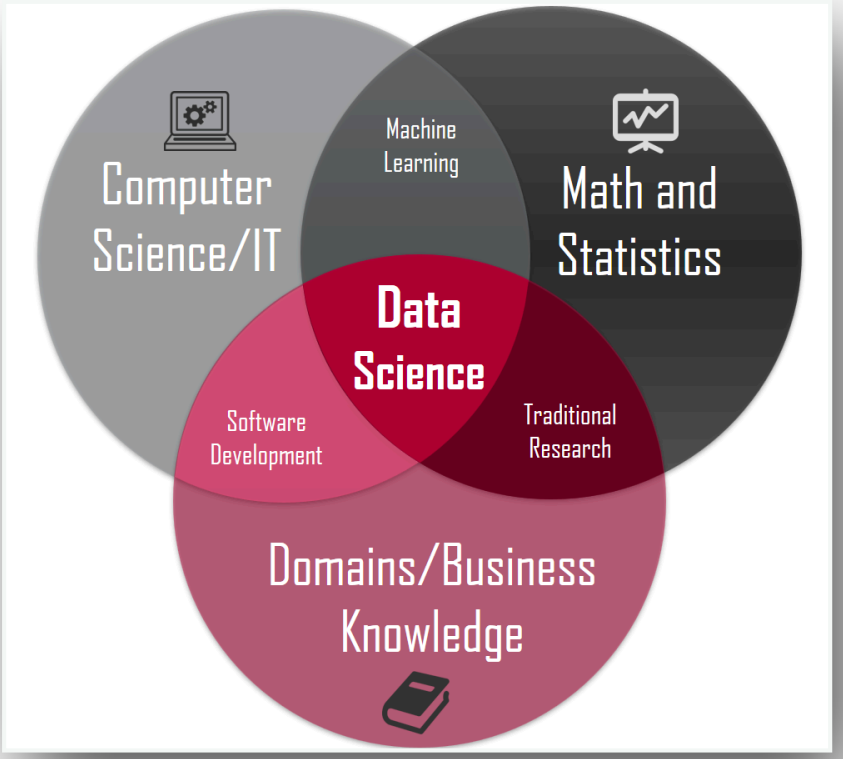
- **Talent**
  - To execute, we need top quality talent:
    - Data Scientists, Behavioral Scientists
    - Data Engineers, Solution Architects
    - Product Managers, Technical Project Managers
  - How do you structure **career path** for folks within the data org?
    - Should they roll up under tech or product or business?

Enterprise Data Strategy + Analytics

- **Talent - Data Science**

- Combines scientific methods & algorithms to extract knowledge & insight from data.

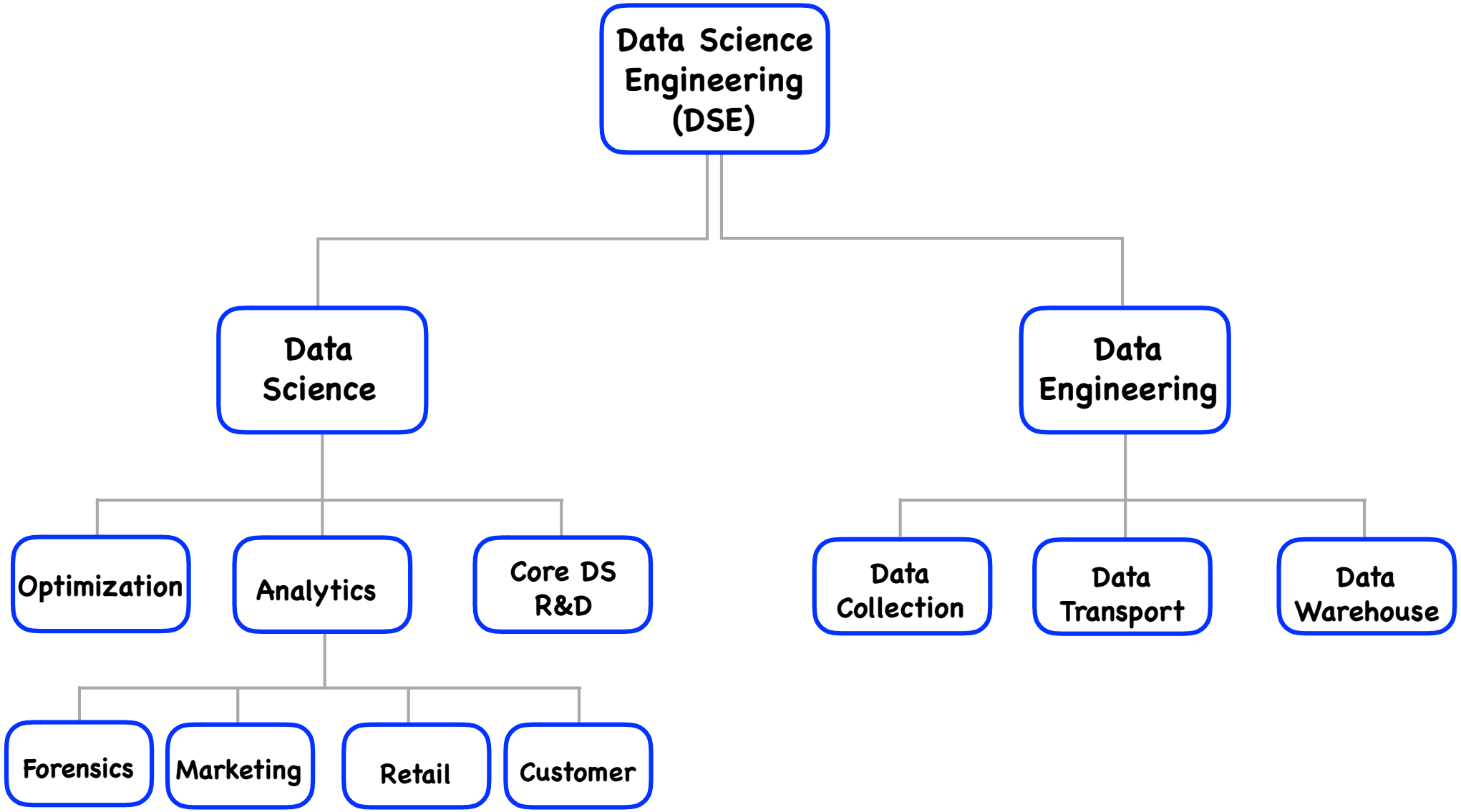
- **Effective data scientists**:
  - high-level **technical** skills
  - build complex **quantitative** algorithms
  - **coding** knowledge
  - Design **experiments**
  - organize and synthesize **large** amounts of data
  - drive **strategy**
  - translate results into **solutions**
  - **communicate** findings



"...The ability to take data — to be able to understand it, to process it, to communicate it — that is going to be a hugely important skill in the next decades..."

—**Hal Varian**
Chief economist at Google & UC Berkeley emeritus professor of information sciences, business & economics

Enterprise Data Strategy + Analytics



Conceptual Talent Organization

Setting Business Analytics Agenda

Setting Business Analytics Agenda

- If data strategy is set up correctly, then business analytics agenda ought to **evolve naturally**.
- Business analytics agenda **CANOT** be **REACTIONARY**.
  - Analytics must be **plugged in** fully.
- In particular, business analytics agenda must address the following at an **enterprise** level:
  - Business problem **formulation**
  - Prescriptive **solutions**
  - Monitoring, reporting & **insights**

Setting Business Analytics Agenda

- **Business Problem Formulation**
  - Business must ask **tough** questions of itself if it wants to remain in business!
    - How do we **innovate**?
    - Who is our **competition**?
    - Who is our **customer**?
    - Why is **churn** rate this high?
    - How do we increase **market share**?
    - How do we improve **customer experience**?...
  - **Business analytics** must provide the **right context** by leveraging data to enable leadership frame these questions appropriately.

Setting Business Analytics Agenda

- **Prescriptive Solutions - Proposals**
  - Once business problems are clearly defined and scoped, business analytics must work with domain experts to **formulate solution proposals**.
  - Solution proposals leverage data to identify areas of **opportunities** to address business challenges.
  - Oftentimes, analytics have a **360 view** of the business so they are able to identify **critical stakeholders** required to execute.
    - **FYI**: Teams must be structured to support this capability

Setting Business Analytics Agenda

- **Prescriptive Solutions - Execution**
  - What solutions are we **implementing** and how? are we ...
    - designing or rolling out targeted **interventions**?
    - building **backend models/algos** to power new features or products?
    - conducting **causal analysis** to identify critical factors for a product?
    - running **experiments** to quantify effects of interventions?
  - Analytics must work closely with stakeholders to provide guidelines around **design**, **datasets**, **methodology**, **outcomes** etc

Setting Business Analytics Agenda

- **Monitoring, Report & Insights**
  - **Define relevant metrics** - if possible please **track everything**!!!
  - Analytics must have **basis** to determine **performance** of new features, roll-outs, products etc.
  - Metrics must be tied to **overall business strategic goals** and not just departmental, group or provincial goals.
  - Business health assessments, evaluations and projections etc must be rooted in insights -
    - analytics is responsible for bringing the right business perspectives to leadership
  - Leveraging metrics for all key decisions drives the org closer to being fully **data driven** - ultimate **business analytics goal**

Setting Business Analytics Agenda

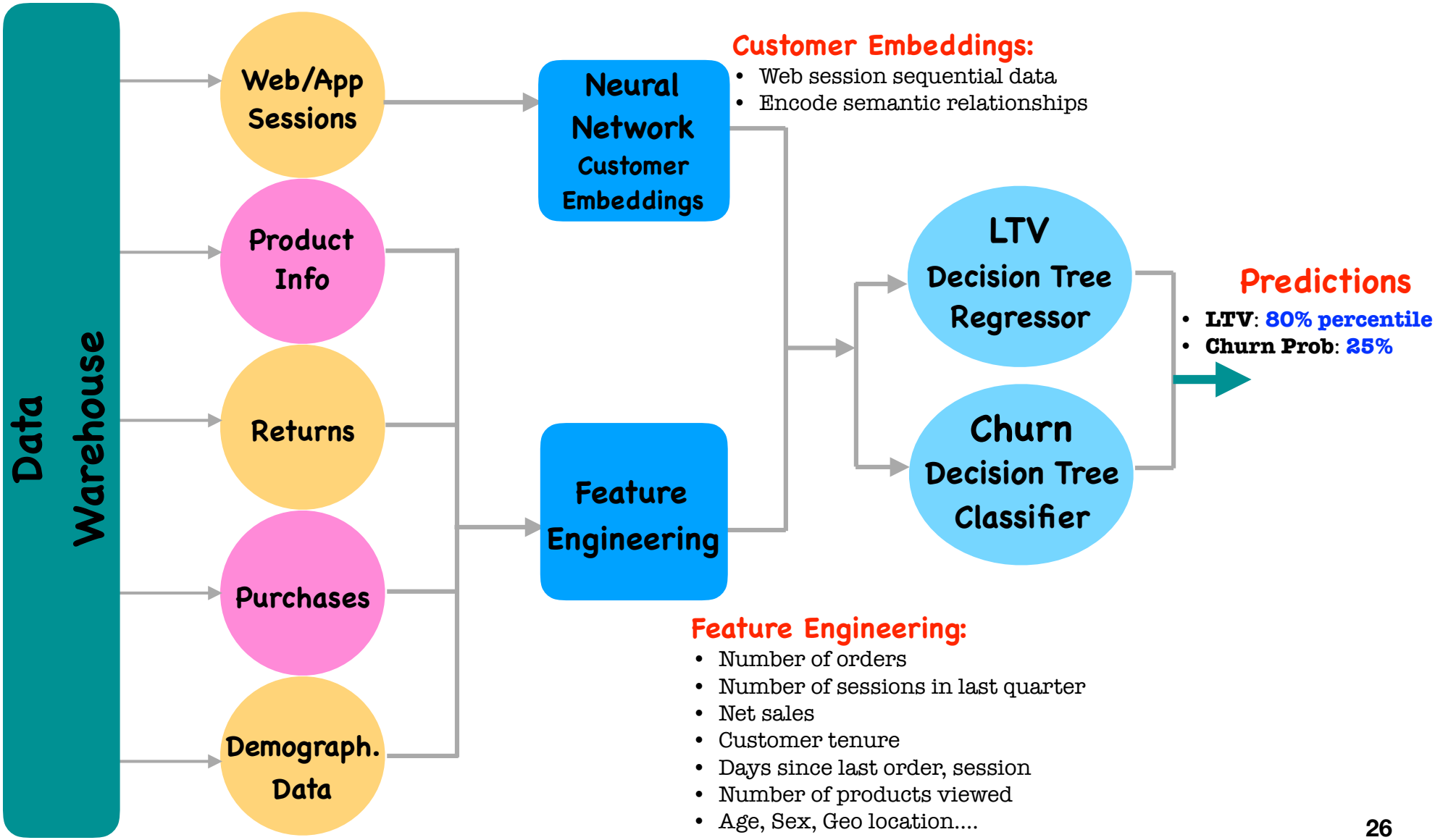
- Ensure your efforts are **not caught up in silos**:
  - **AI adoption** and **projects** need to be:
    - cross-functional
    - across leadership levels
    - collaborative
  - **AI** projects require a **synthesis** of:
    - AI expertise
    - domain knowledge
    - business acumen
    - corporate strategy and vision

AI StoryBoarding - Strategic Use Cases

Use Case I:
Customer Lifetime Value

- What is **CLV?**
  - The value of a customer to your business over the entire length of your relationship with the customer
    - e.g sales net of returns over a time period, say one year.
- What is **Churn Probability?**
  - Likelihood of a customer not buying from you over a time period
- **CLV** & **Churn** are critical marketing metrics particularly so for e-Commerce businesses
  - **Customer relationship management**: high-valued customers retention
  - **Marketing strategy**: budgeting, targeted campaigns

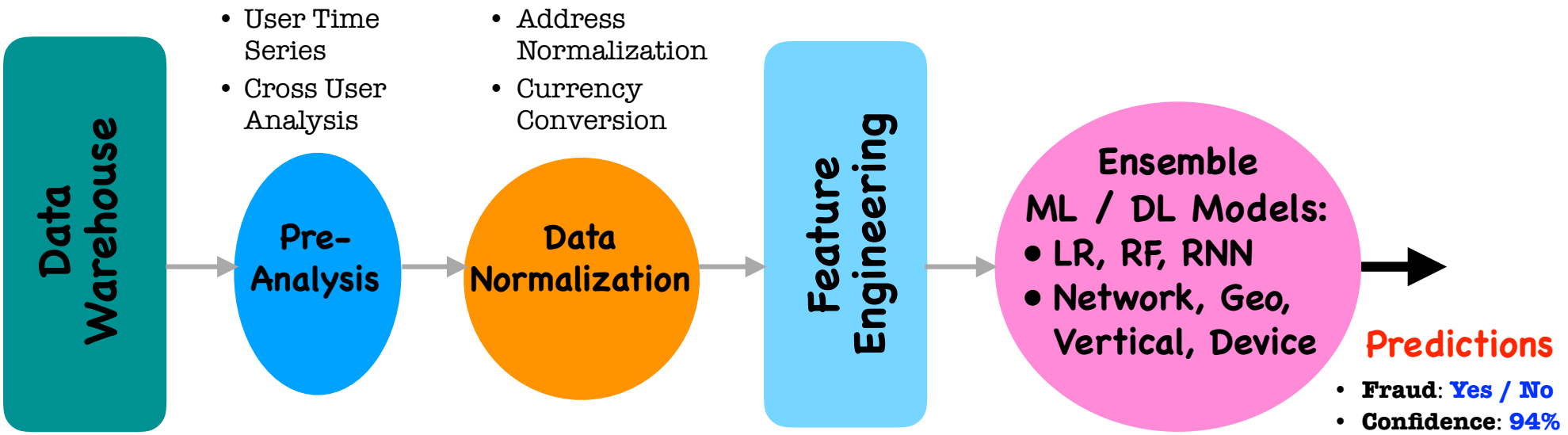
Use Case I:
Customer Lifetime Value



Use Case II: Fraud Detection

- **Fraud** detection is **challenging**!
  - Fraudulent transactions are **rare**
  - small sample but potentially **significant** losses
  - Fraudsters are **crafty** and fast **evolving**
- Most companies use **rule-based** systems
  - Effective if pattern is already **known**
  - Canot **uncover unknown** schemes
  - Canot **adapt** to **new** fraud patterns or techniques
- ML models can **learn**, **adapt** and **uncover** emerging fraud patterns

Use Case II: Fraud Detection



Data Types:

- User identity
- Behavioral patterns
- Locational data
- Device & network
- Business Unique Data
- Business Decisions
- Transactions
- Third Party Data

Feature Types:

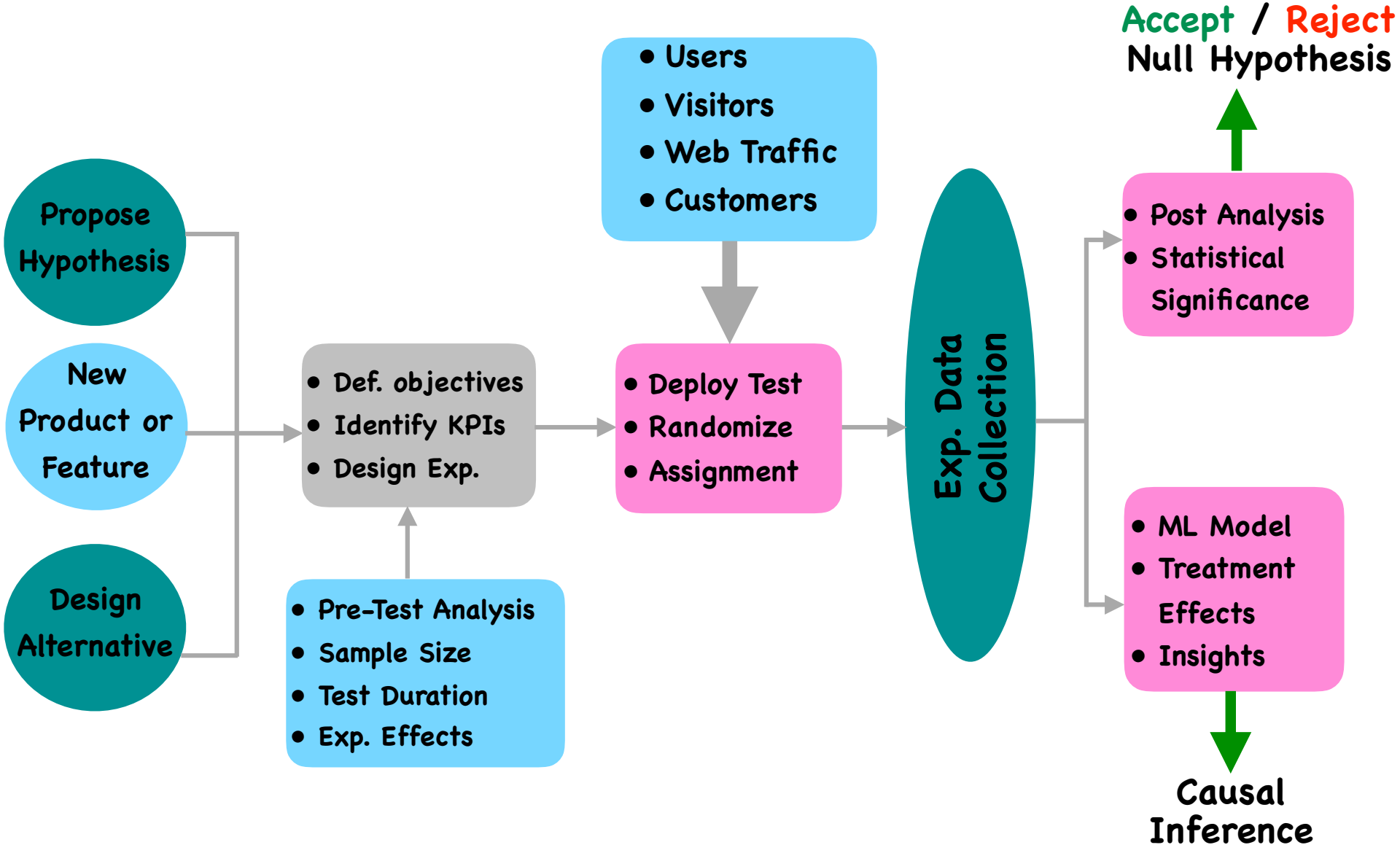
- Event feature
- Geo feature
- Temporal feature
- Identity feature
- Velocity feature
- Behavior feature

Feature Types are categorial & continuous.

Use Case III:
A/B Testing

- **Customer addition** & **satisfaction** is critical for business success:
  - How do you **attract new** customers?
  - How do you **retain existing** customers?
  - How do you **roll out new** products or experiences?
- **Experimentation** - how big tech companies improve **customer experience**.
  - Test **hypothesis**
  - Determine **causality**
  - Launch, debug, measure and monitor effects of:
    - **enhancements,** product **features,** marketing **campaigns**
    - backend **models** e.g. search, recommendations, pricing algorithms

Use Case III: A/B Testing




Thank
You