## **Workshop Assignment: Architecting for the AI Era – Designing Modern Data & Information Architecture**

### **🎯 Objective**

To evaluate, design, and present a modern data architecture that supports AI readiness, data agility, and value delivery across business units. Participants will use real-world scenarios and the six foundational shifts to reimagine their enterprise's architecture, focusing on both data and information architecture integration.

### **📚 Pre-Reading**

* *Modern Data Architecture and AI* by Daniel Lambert
* TOGAF definitions of Data & Information Architecture
* Concepts from McKinsey’s six shifts in data architecture

### **🛠 Assignment Structure**

#### **Part 1: Diagnostic (Team Exercise – 45 min)**

**Objective:** Evaluate your current-state data architecture.

**Instructions:**

1. In small groups (3–5), map your enterprise’s current data architecture components (data warehouse, lake, marts, integration tools, etc.).
2. Identify the **gaps** using the following prompts:  
   * Where is data agility compromised?
   * How accessible is data to AI model developers?
   * Is there a clear separation between data platforms and architecture?
   * What percentage of “data” is actually usable “information”?

**Deliverable:** A whiteboard (digital or physical) visualization of your current architecture and 3 major obstacles to AI enablement.

#### **Part 2: Strategy Rebuild – Applying the Six Foundational Shifts (Solo or Pair – 60 min)**

**Objective:** Redesign your data architecture to align with modern principles.

**Instructions:** Use the following shifts as redesign levers:

* From on-prem to cloud-based
* From batch to real-time
* From commercial to modular
* From point-to-point to decoupled
* From enterprise warehouse to domain-driven
* From rigid to flexible schemas

**Task:** Draw a **“future-state” high-level architecture** that supports AI and analytics innovation using these shifts. Highlight:

* Key platforms (e.g., Snowflake, BigQuery, Databricks)
* Stream/data flow (real-time, batch)
* API/data product layers
* Security/governance overlays

**Bonus:** Identify which parts can be implemented without replacing legacy systems.

**Deliverable:** Future-state architecture diagram with rationale (bullet list or short narrative).

#### **Part 3: Information Architecture Challenge (Solo – 30 min)**

**Objective:** Identify core *Information Concepts* supporting a customer-facing AI use case.

**Instructions:** Pick **one AI use case** such as:

* Personalized product recommendations
* Intelligent chatbot
* Predictive maintenance
* Fraud detection

**Then:**

1. List 5–7 core information concepts (e.g., customer profile, usage logs, ticket history).
2. For each concept, answer:  
   * Who creates/uses it?
   * Where and when is it accessed?
   * What transformations are applied before it's usable?
   * What quality, privacy, or governance needs exist?

**Deliverable:** A table of **information concepts** with attributes and usage flow. Bonus if mapped to business capabilities.

#### **Part 4: Group Presentation & Discussion (30 min/team)**

**Objective:** Share redesigned data & information architecture and discuss trade-offs.

Each team presents:

* Current-state weaknesses
* Future-state design
* Key data/information enablers for AI
* Implementation risks and change priorities

Facilitator or peer Q&A follows each team.

**Assignment Objective:** Redesign your organization’s data architecture to enable scalable, flexible, and AI-ready capabilities by applying the principles of Modern Data Architecture (MDA) and Information Architecture (IA). The goal is to create a future-ready blueprint for business agility, AI adoption, and information clarity.

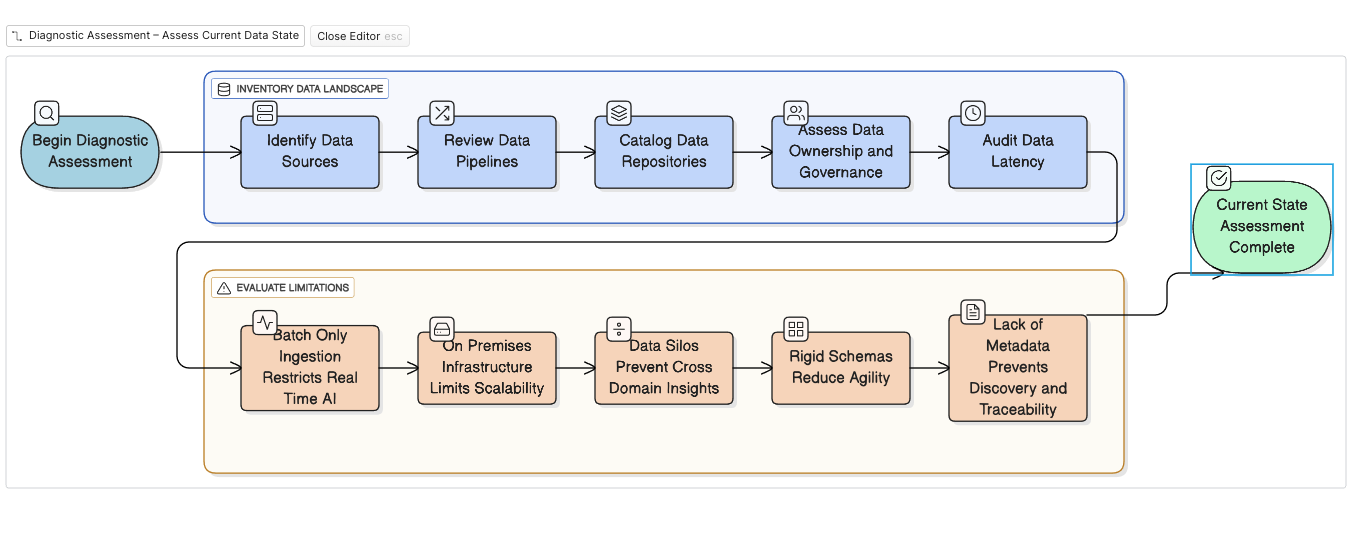
**Section 1: Diagnostic – Assess Current State**

**1.1 Inventory Existing Data Landscape**

* Identify data sources: operational databases, CRMs, ERPs, spreadsheets, legacy systems.
* Review data pipelines: ETL jobs, data integration workflows, streaming systems.
* Catalog data warehouses, marts, and lakes in use.
* Assess data ownership and governance practices.
* Audit latency: time from data capture to insight.

**1.2 Evaluate Limitations**

* Batch-only ingestion restricts real-time AI capabilities.
* On-prem infrastructure limits scalability.
* Data silos prevent cross-domain insights.
* Rigid schemas reduce agility in model development.
* Lack of metadata prevents intelligent discovery and traceability.



**Section 2: Design Modern Data Architecture (MDA)**

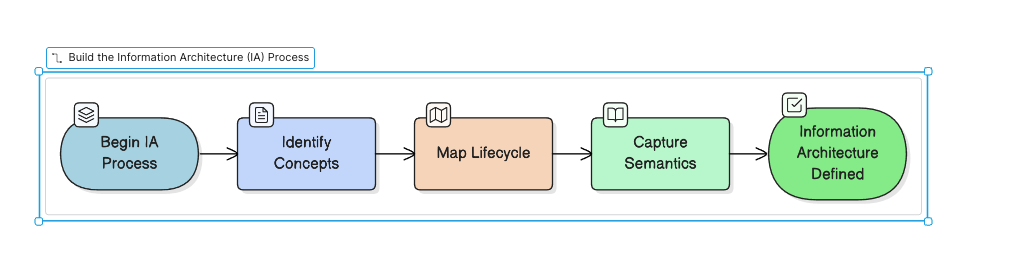
**Apply the Six Foundational Shifts:**

1. **From On-Premises to Cloud-Based Platforms**
   * Migrate legacy systems to AWS Redshift, Azure Synapse, Snowflake, etc.
   * Enable elastic scaling, pay-per-use cost model.
2. **From Batch to Real-Time Processing**
   * Introduce Apache Kafka, Amazon Kinesis, or Azure Event Hubs for streaming.
   * Enable real-time fraud detection, alerts, and recommendations.
3. **From Monolithic to Modular Tools**
   * Replace closed suites with composable services: dbt, Airbyte, Looker, etc.
   * Focus on best-of-breed pipelines, observability, and CI/CD practices.
4. **From Point-to-Point to Decoupled Access**
   * Deploy Data Mesh or Data Fabric architectures.
   * Provide API-driven and event-based access to domain data products.
5. **From Enterprise Warehouse to Domain-Based Architecture**
   * Assign ownership of data products to business domains (sales, HR, ops).
   * Create reusable, governed datasets as a service.
6. **From Rigid Models to Flexible Schemas**
   * Embrace schema-on-read in data lakes.
   * Use flexible storage like Delta Lake, Iceberg, or Hudi.
   * Support semi-structured data: JSON, Parquet, Avro.

**Section 3: Build the Information Architecture (IA)**

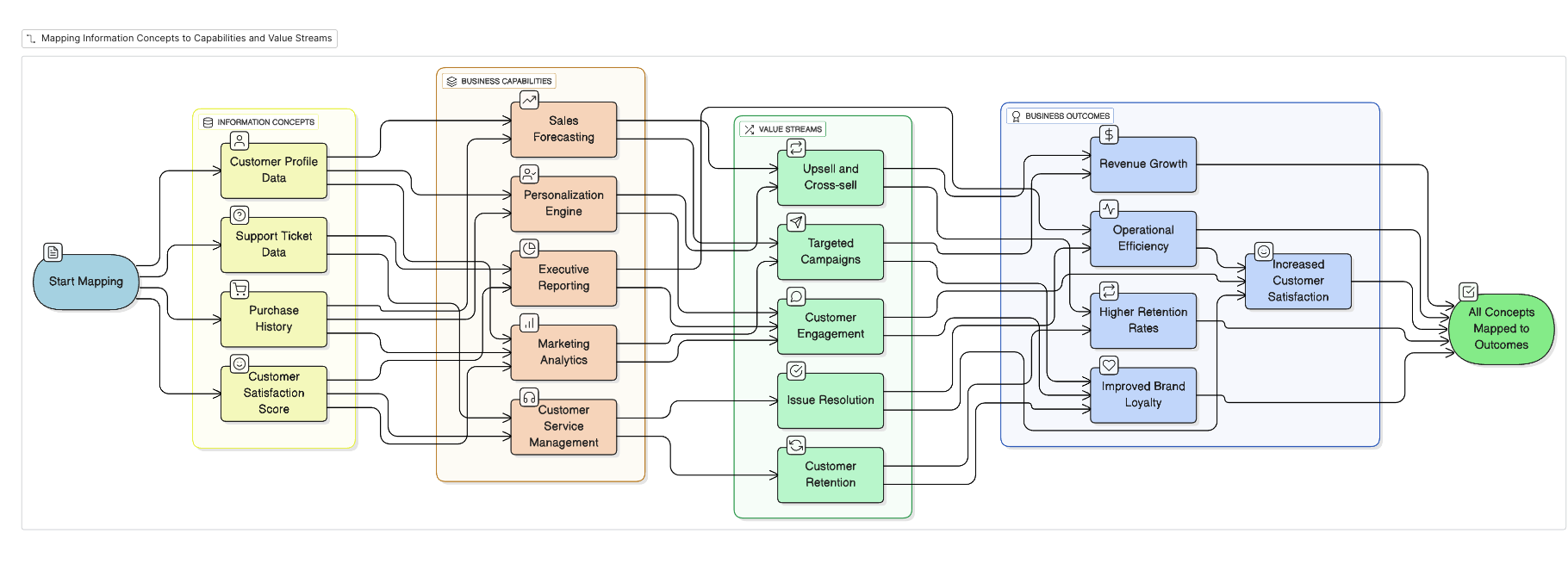
**3.1 Identify Information Concepts**

* Define key business concepts (e.g., customer, order, product, asset).
* Map where they are created, modified, used (across systems).
* Capture semantics: definitions, lineage, stewardship.



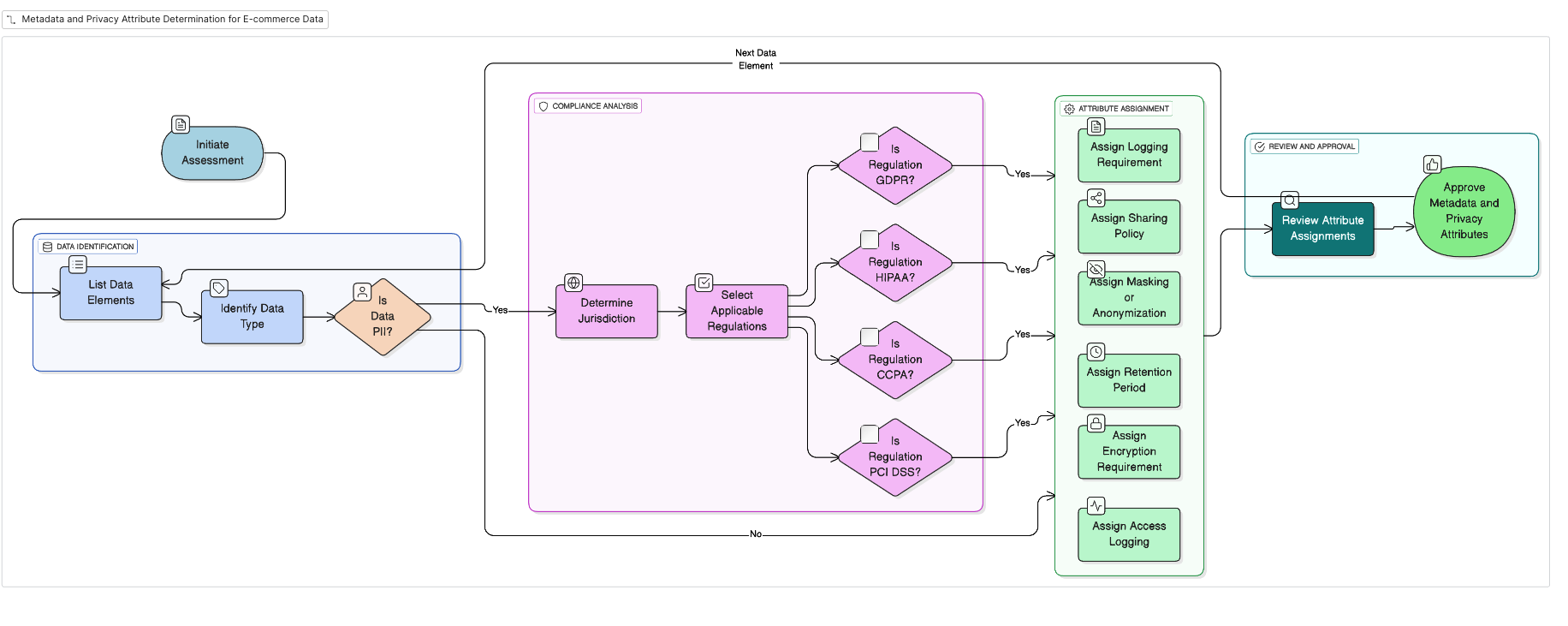
**3.2 Map to Capabilities and Value Streams**

* Link each information concept to business capabilities and outcomes.
* Example: “Customer Satisfaction Score” used in marketing, service, and executive dashboards.



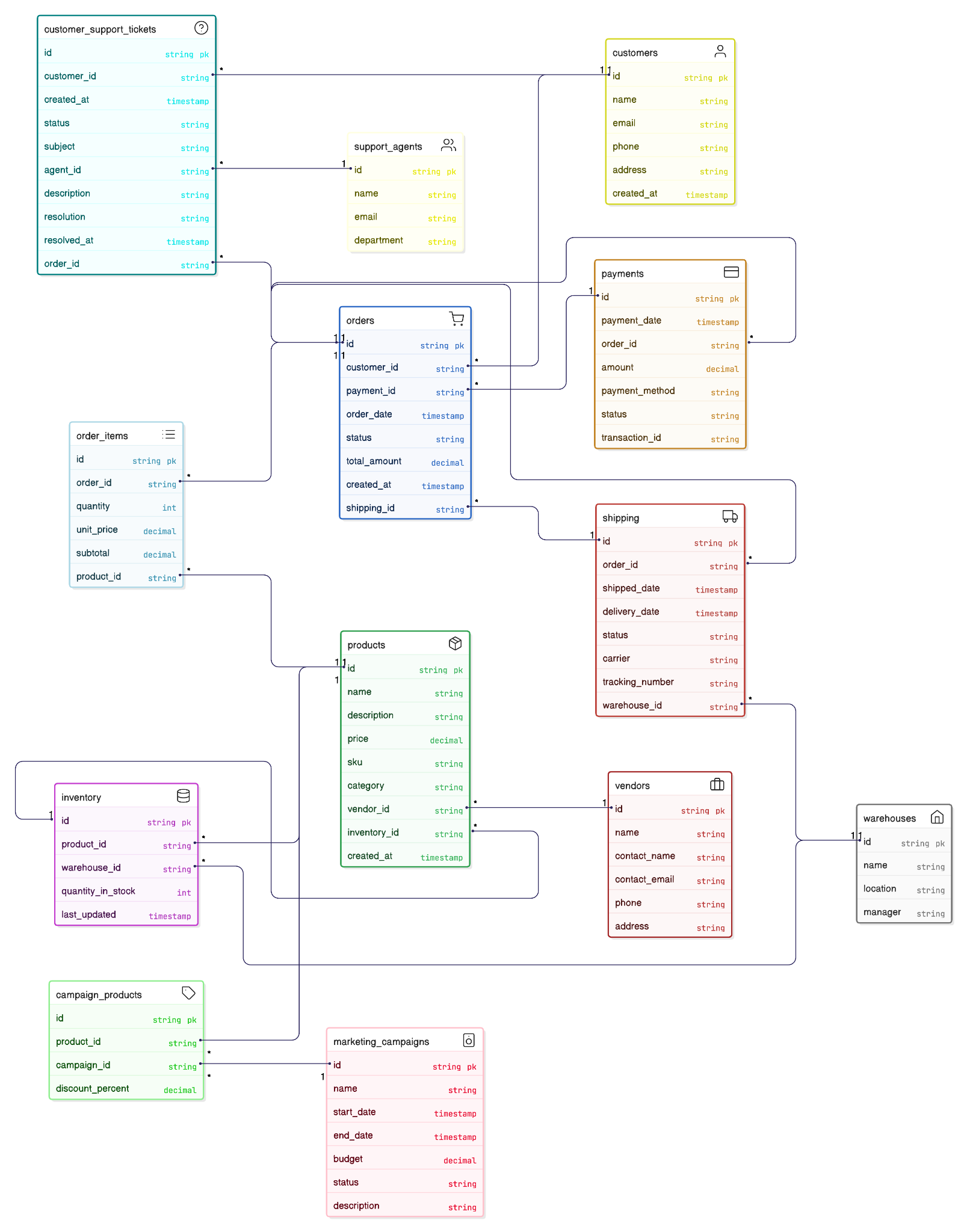
**3.3 Determine Metadata and Privacy Attributes**

* What info must be logged, shared, masked, or retained?
* Identify PII, compliance constraints (e.g., GDPR, HIPAA).



**3.4 Build Visual Model**

* Construct Concept Maps, Information Flow Diagrams, and Glossaries.
* Tools: Lucidchart, Miro, ER/Studio, or ArchiMate.

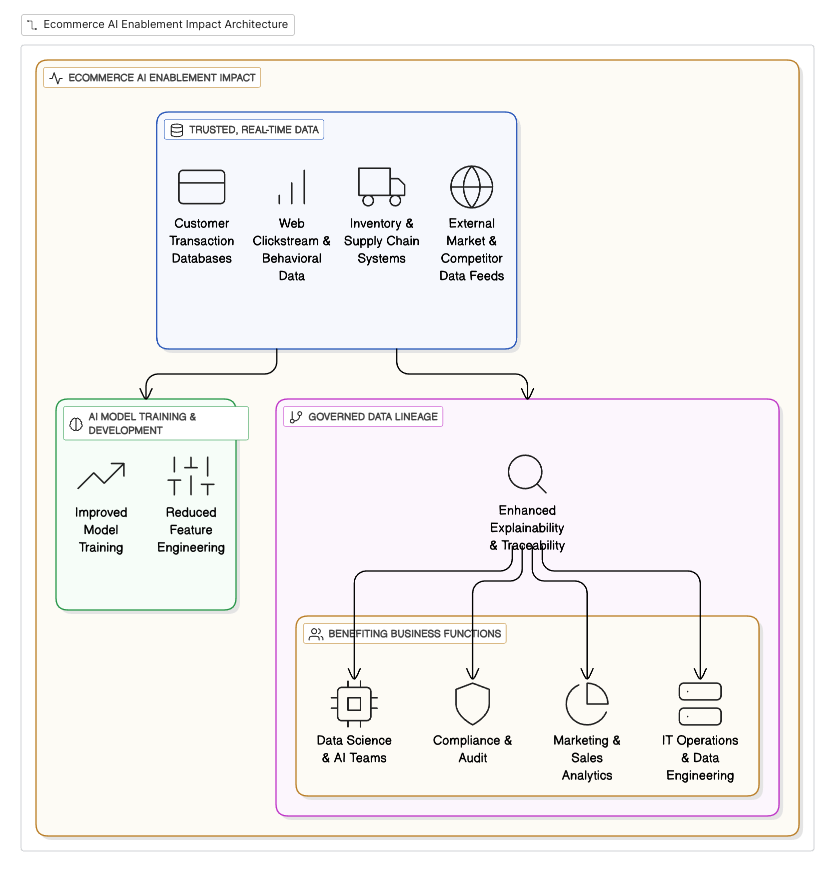




**Section 4: Executive-Level Architecture Summary**

**4.1 AI Enablement Impact**

* Improved AI model training via access to trusted, real-time data.
* Reduced feature engineering effort due to clean, well-modeled data.
* Enhanced explainability and traceability from governed lineage.



**4.2 Organizational Impact**

* Faster insights across business units.
* Accelerated digital product innovation.
* Tighter governance and reduced data duplication.

**4.3 Risk Mitigation**

* Resilience to market shifts through agility.
* Reduced compliance risk from secure, well-managed information flows.

**Section 5: Implementation Roadmap (3 Phases)**

**Phase 1 – Foundation (0-3 Months)**

* Set up cross-functional data strategy team.
* Select modern data platform (e.g., Snowflake, Databricks).
* Define and prioritize information concepts.
* Begin pilot: streaming ingestion + AI-ready dataset.

**Phase 2 – Expansion (3-9 Months)**

* Migrate domain data products to cloud and mesh-based governance.
* Enable API and event-driven data services.
* Embed real-time data feeds into AI systems.

**Phase 3 – Optimization (9+ Months)**

* Standardize metadata management and information cataloging.
* Integrate data quality, observability, and AI bias detection tools.
* Extend platform for generative AI, GraphRAG, and predictive services.

**Success Metrics**

* Time-to-insight reduced by X%
* Model accuracy improved by Y%
* Cost per insight reduced by Z%
* Data reuse increased across business domains
* Data literacy index and catalog adoption rate

**Conclusion:** By combining modern data architecture and robust information architecture, Enterprise Architects enable their organizations to harness the full power of AI—responsibly, scalably, and strategically. This assignment gives you the structure to deliver lasting value across business and technology domains.

Let me know if you'd like industry-specific examples or visual slide templates to go with this solution.