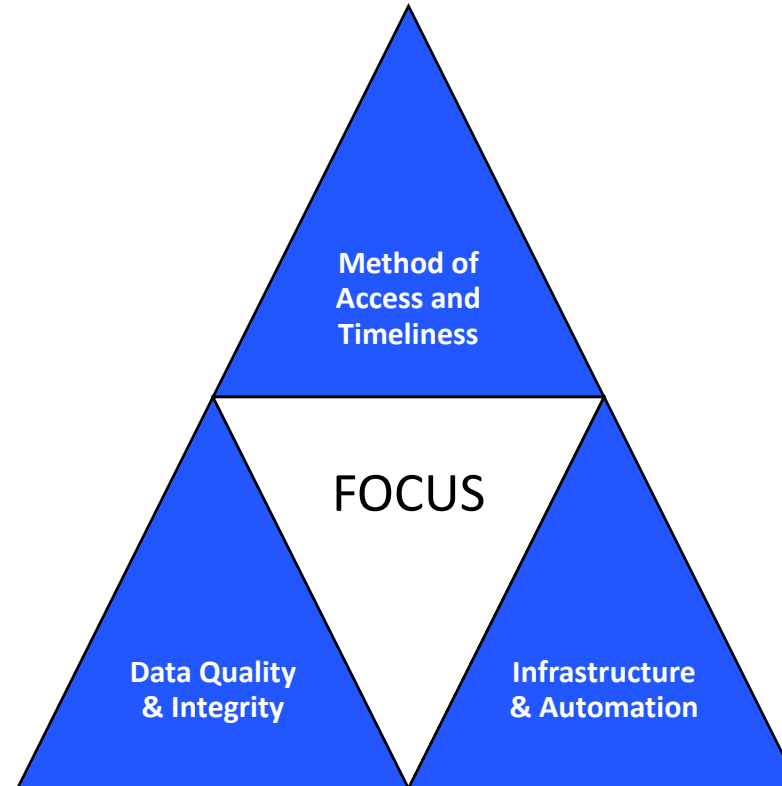BLUESHIFT INNOVATIONS INC.

Modern Data Platform

Blueshift™

"Trifecta" of common data modernization themes

All 3 areas represent the most challenging aspects of todays systems, and it is recommended to addressed all 3 simultaneously. Focusing only on any subset will be suboptimal.

- For example: Focusing only on **Data Quality & Integrity** so that it has a high-degree of credibility and trust may be an improvement, but the Data would still be very **difficult to access** and not on a 'rock-solid, scalable, reliable' **infrastructure**.

OR

- For example: Focusing only on a 'rock-solid, scalable, and reliable' **infrastructure** and an improved **method of access** will improve the user experience, but to inferior and unreliable **data**.

Modern Data Platform

A managed cloud-native data platform that delivers speed-to-market, confidence, and agility for data modernization problems at a lower total cost of ownership (TCO).

Metadata Driven

Decouple data model from underlying architecture to create agility. Traditional data warehouses require static models that cannot evolve with your business.

Data Lakehouse

Architecture pattern across all stages of the data lifecycle – from ingestion to visualization. Traditional data solutions are purpose-built for a specific stage

Data Fabric

Simplify data access by automating data discovery, enabling self-service consumption and governance despite where data resides in the enterprise

Data Mesh

Handle vast, distributed data consumers and views with domain-specific handling in data pipelines - unlike the monolithic architecture of legacy systems.

Data Virtualization

Access data at the source without moving data, accelerating time to value with faster and more accurate queries.

Data Governance

Processes, policies, and standards that protect data & maintain integrity, and make it available for consumers - with lineage and explainability of how and why data has moved over time.
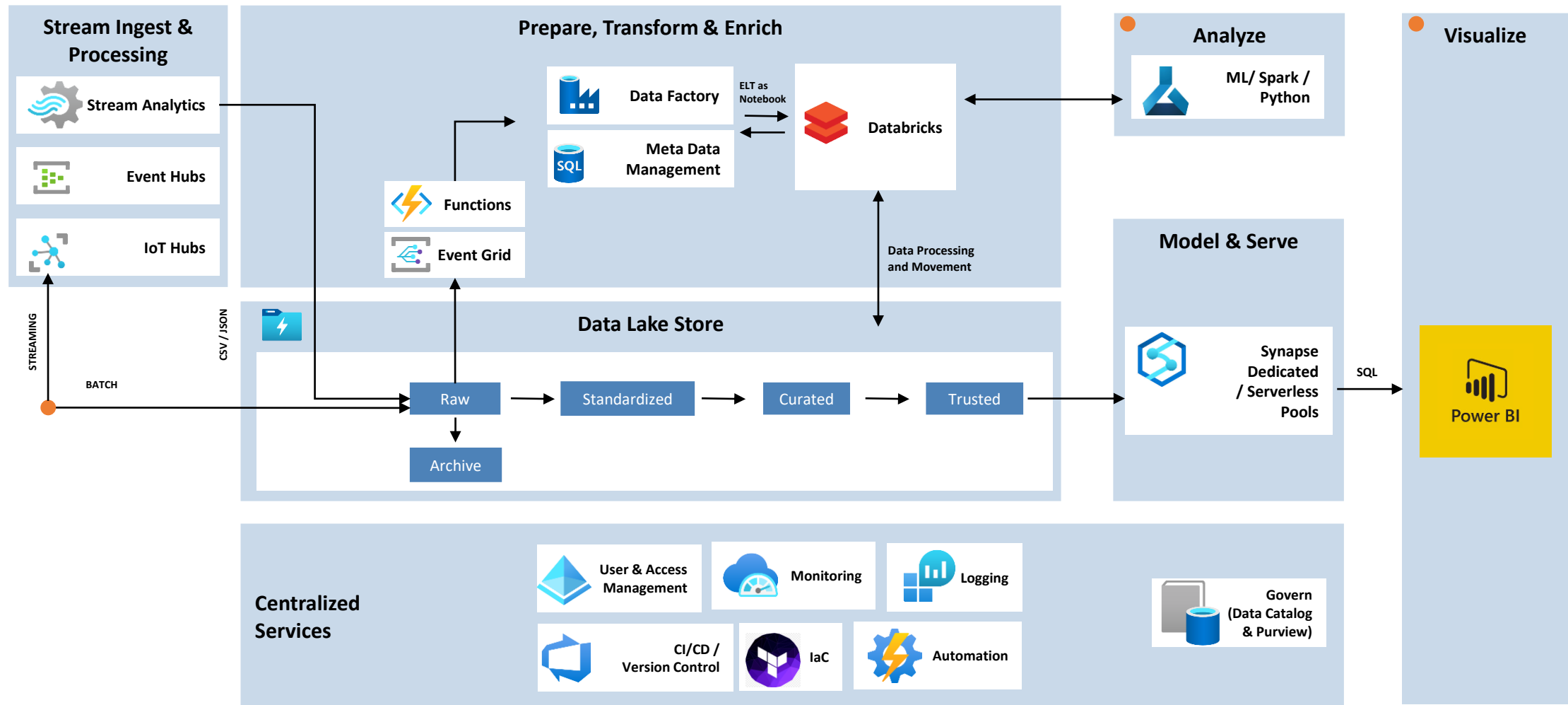
Modern Data Platform based on a Lakehouse Pattern
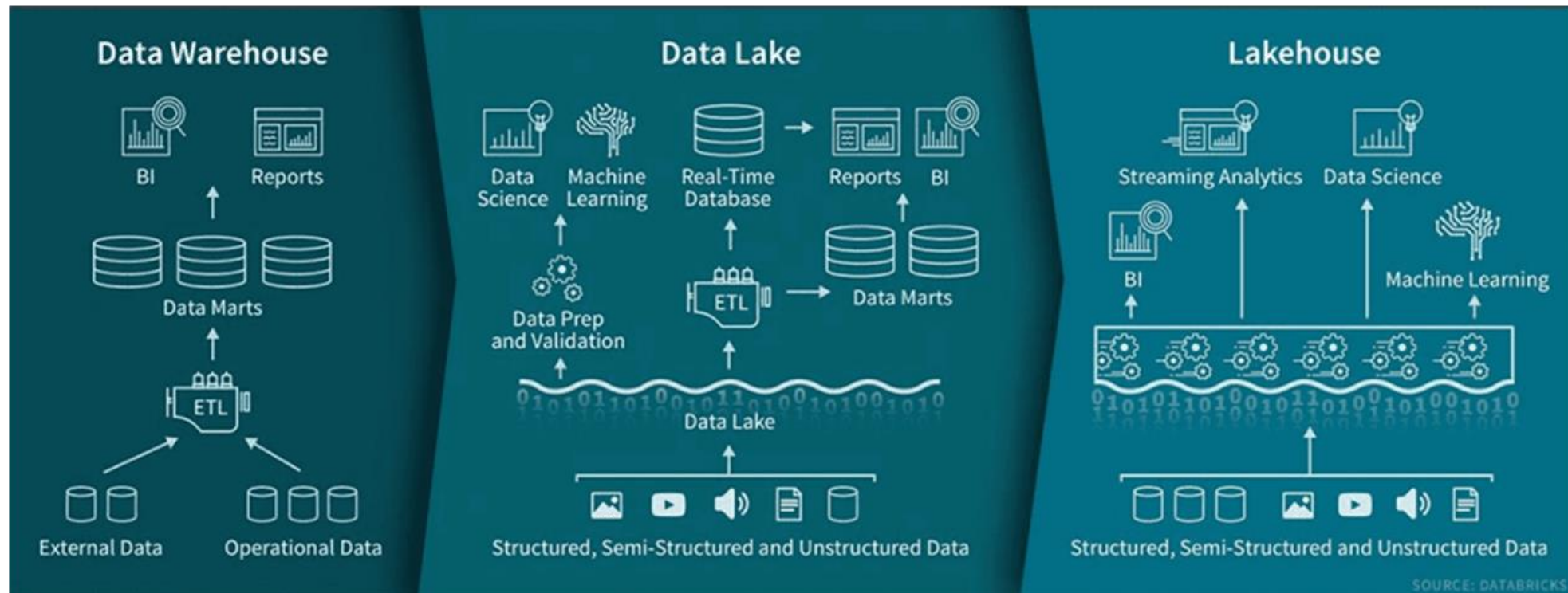
Modern Data Platform based on a Lakehouse Pattern

ARCHITECTURE MODEL	BUSINESS PROCESSES REQUIRED				
	INGEST	**EXPLORE & PREPARE**	**TRANSFORM & ENRICH**	**MODEL & SERVE**	**VISUALIZE**
Data Warehouse	Requires integration with 3rd party technology via ETL drivers	Requires integration with 3rd party technology via ETL tools, and continuous engineering to maintain. ETL *all* data, during which each ETL step risks incurring failures that reduce data quality. Infrastructure is tightly coupled, creating scalability issues with cost profile for peak user loads.		**BEST** Purpose-built for **MODEL & SERVE**. • Focused on organize, normalize, integrate, and quality of data. • Designed for *schema-on-write*. • Supports *structured data* only.	Requires integration with 3rd party BI tools
Data Lake	**BEST** Purpose-built for **INGEST**. • Focused on ingesting large volumes of raw data, in variety of formats, on cheap storage. • Designed for *schema-on-read*. • Supports *Structured, Semi-structured, and Unstructured* data only.	Requires integration with 3rd party technology via ETL tools, and continuous engineering to maintain. ETL *small subset* of data, and still no ACID (atomicity, consistency, isolation, durability) transaction support to ensure quality and reliability. Problem of data quality and governance is punted downstream to the warehouse.		Requires integration with 3rd party Data warehouse system Data here is *stale* relative to the Lake, from which can takes *days+* to load and re-load – two-tier architecture. These warehouses are not designed to solve for quality or governance, and only contain a small subset of data in the lake.	Requires integration with 3rd party BI tools
Lakehouse	**BEST** Purpose-built as a **single, unified and consistent end-to-end** platform. Optimized for performance, cost, reliability and scalability to deliver capability at each stage of the business process. Support for *directly-accessible storage* (like traditional Data Warehouses) for ACID transaction, data versioning, auditing, indexing, caching, and query optimization. Supports advanced analytics via non-SQL code: Respond to business-oriented questions, not just typical BI queries.				

We've done this before. Modern Data Platforms at Scale.

Azure Landing Zone follow Azure Cloud Adoption Framework best practices
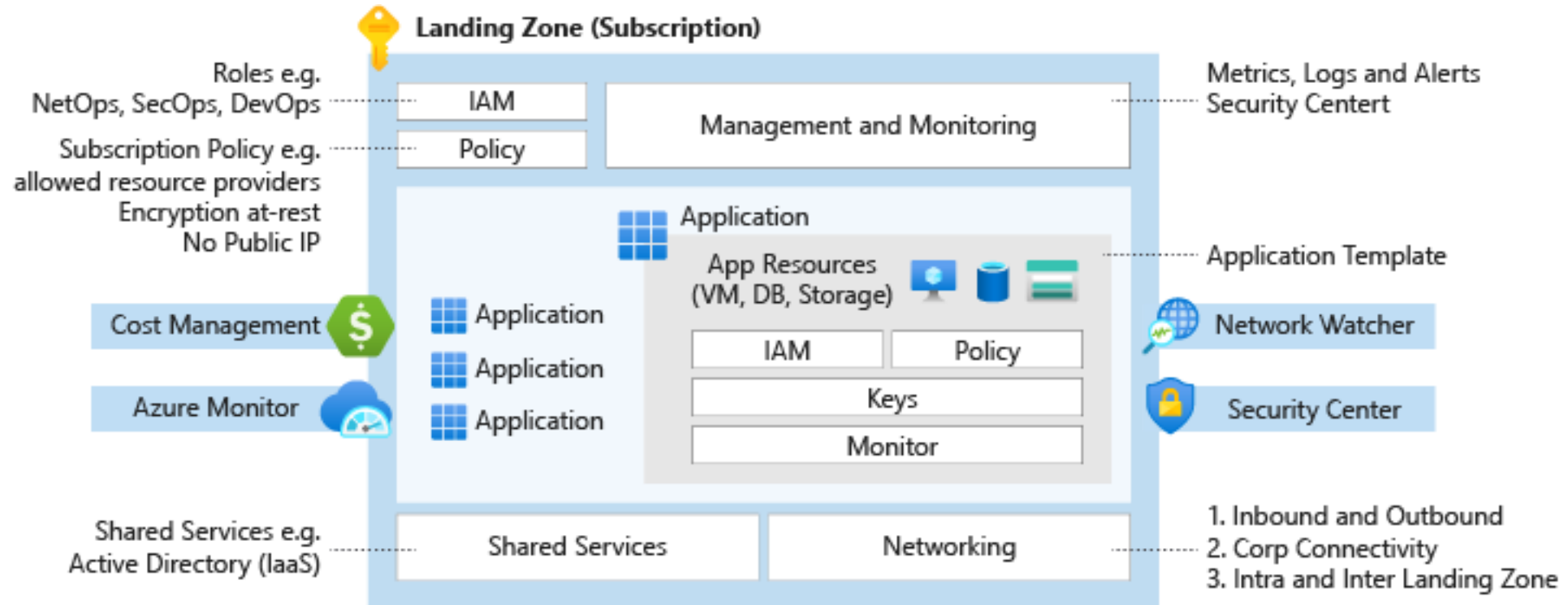
Initial Timeline

We work through series of MVPs to configure the platform.

MVP 1

MVP 2

MVP 3

Final Product

Managed Services
On-going

Blueshift