

Del Origen a la Visión: El Viaje de los Datos

Meibelyn Robles
Leda Araya



SJO DIGITAL 2025



AGRADECIMIENTO ESPECIAL A NUESTROS PATROCINADORES



Primus Data



BRUNA
GROUP

NESCAFÉ®

TEC | Tecnológico
de Costa Rica



advancelearning



AGENDA


Flujo de un proceso de Data Analitica

¿Se requiere Big Data?

Arquitectura de Bus

Arquitecturas Modernas

Demos



Procesos de transformacion de Datos



EXTRACCION

RECOPIACIÓN DE DATOS DESDE DIVERSAS FUENTES Y CARGA EN EL DATA LAKE.



LIMPIEZA Y PREPARACION

PROCESO DE LIMPIAR DATOS CORRUPTOS O INCOMPLETOS Y ESTRUCTURARLOS PARA SU ANÁLISIS POSTERIOR.



TRANSFORMACION

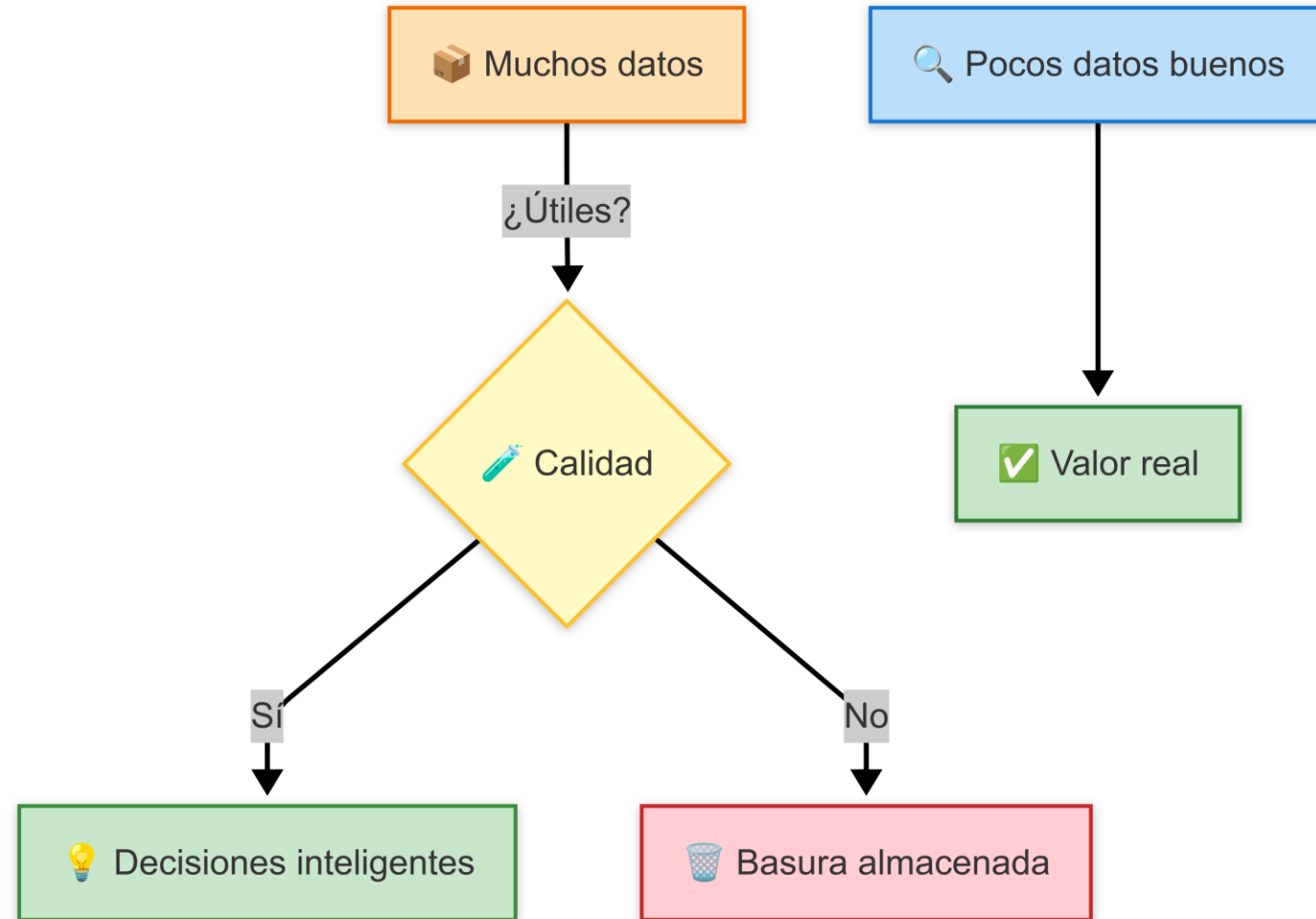
APLICACIÓN DE REGLAS DE NEGOCIO Y PROCESOS DE TRANSFORMACIÓN PARA PREPARAR LOS DATOS PARA SU VISUALIZACIÓN.



ENRIQUECIMIENTO DE LOS DATOS

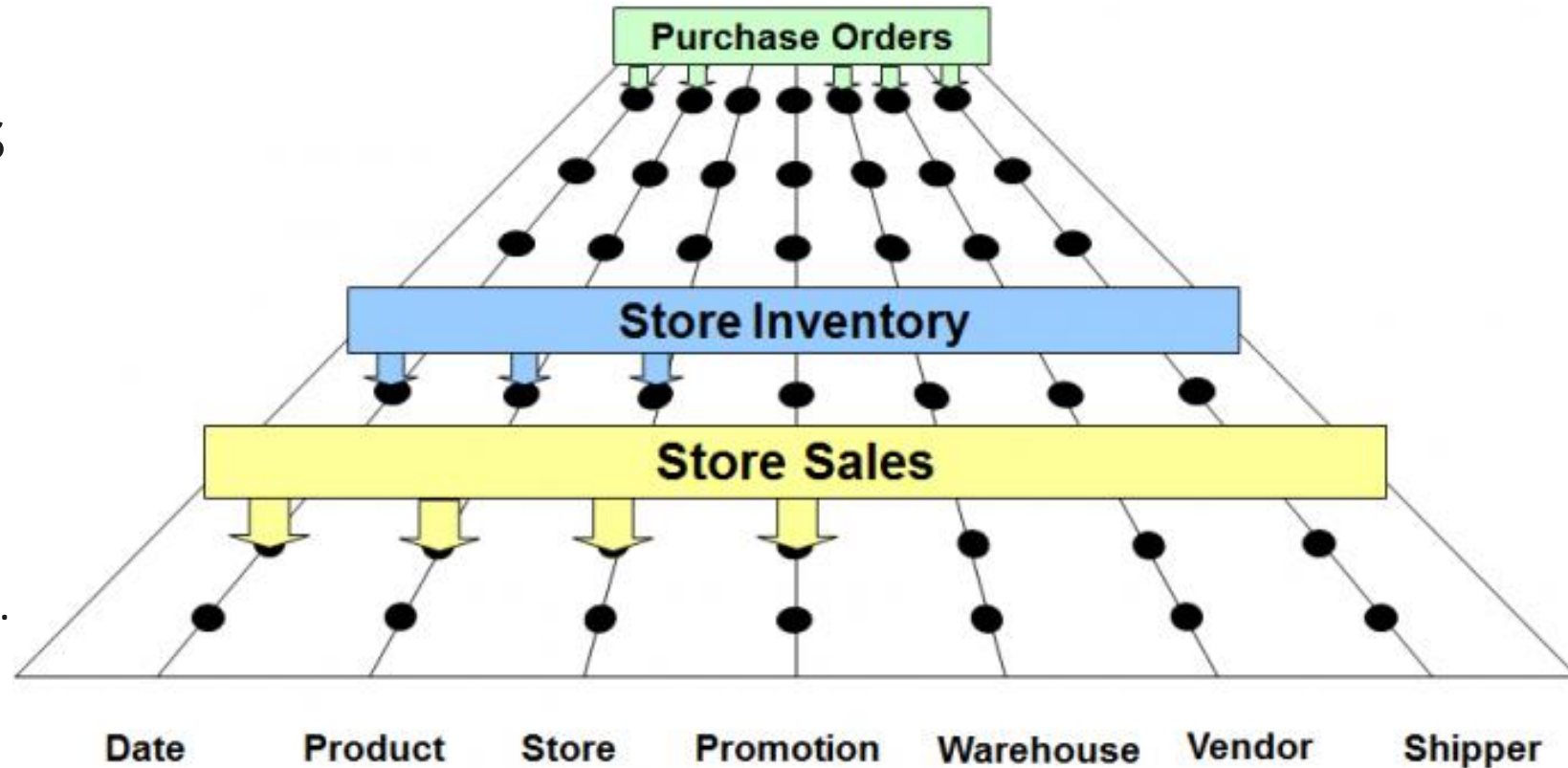
COMBINACIÓN DE DATOS DE DIFERENTES FUENTES PARA OBTENER UNA VISIÓN MÁS COMPLETA.

¿Necesitamos big data?



Arquitectura de bus

Introducida por Kimball,
dimensiones maestras y
comunes a procesos de negocio.



<https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/kimball-data-warehouse-bus-architecture/>

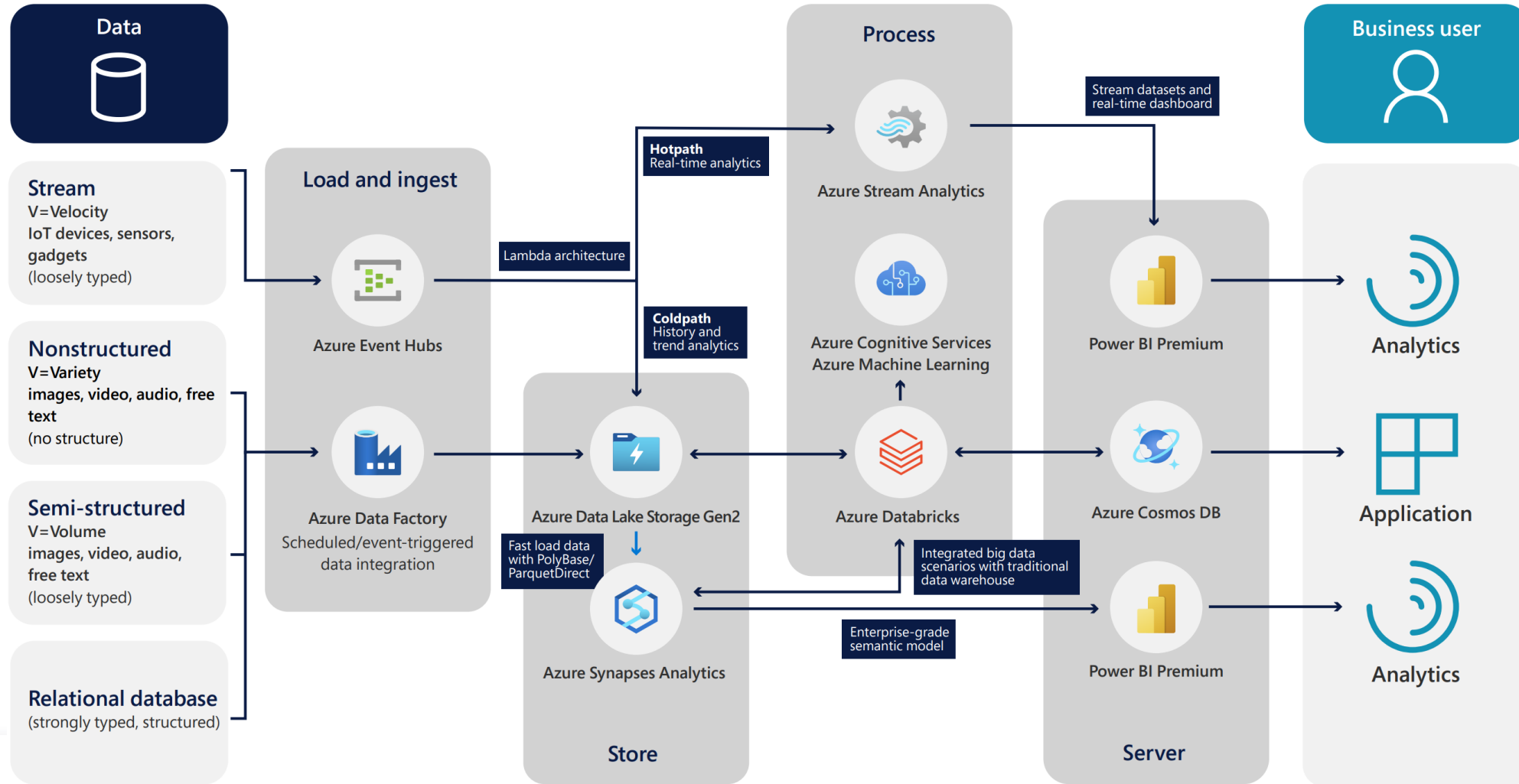
COMMON DIMENSIONS

BUSINESS PROCESSES	Date	Product	Warehouse	Store	Promotion	Customer	Employee
Issue Purchase Orders	X	X	X				
Receive Warehouse Deliveries	X	X	X				X
Warehouse Inventory	X	X	X				
Receive Store Deliveries	X	X	X	X			X
Store Inventory	X	X		X			
Retail Sales	X	X		X	X	X	X
Retail Sales Forecast	X	X		X			
Retail Promotion Tracking	X	X		X	X		
Customer Returns	X	X		X	X	X	X
Returns to Vendor	X	X		X			X
Frequent Shopper Sign-Ups	X			X		X	X

Herramienta clave de diseño: un plato arquitectónico guía de la integración de datos empresarial

<https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/kimball-data-warehouse-bus-architecture/>

Modern Data Platform



Pre-Fabric



Característica

Descripción

Desacoplamiento de almacenamiento y computación

Permite escalar cada uno según necesidad. Ej: separar OneLake (almacenamiento) de Power BI o Spark (procesamiento).

Arquitectura abierta e híbrida

Soporte para datos estructurados, semiestructurados y no estructurados.

Pipeline de datos automatizado

Uso de orquestadores como Data Factory o Synapse Pipelines.

Gobernanza centralizada

Catálogos, políticas de acceso, linaje de datos con herramientas como Purview.

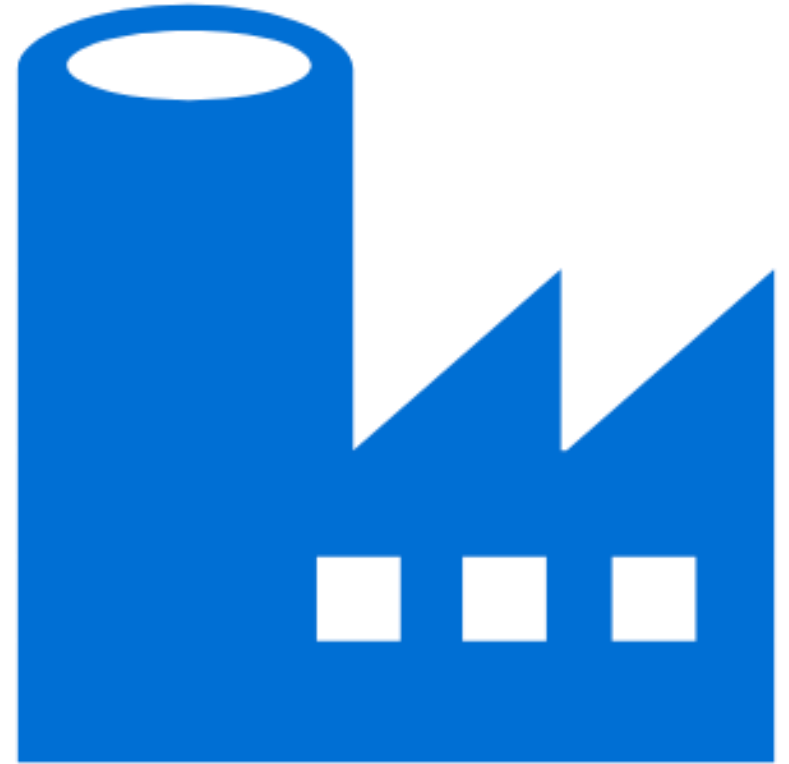
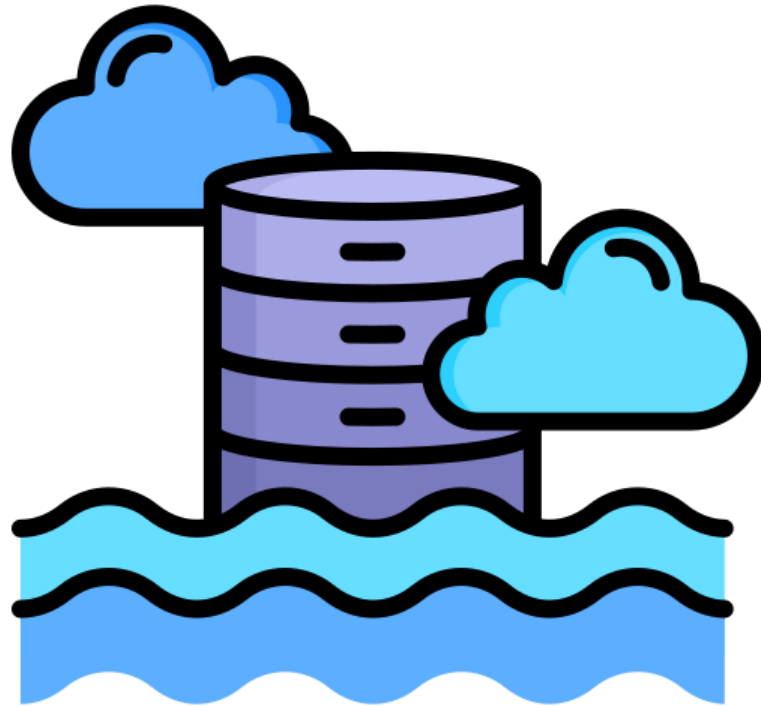
Interoperabilidad

Integración con herramientas de ciencia de datos, BI, machine learning.

Tiempo real y batch

Capacidad de procesar datos históricos y en tiempo real.

DEMO



DEMO Fabric

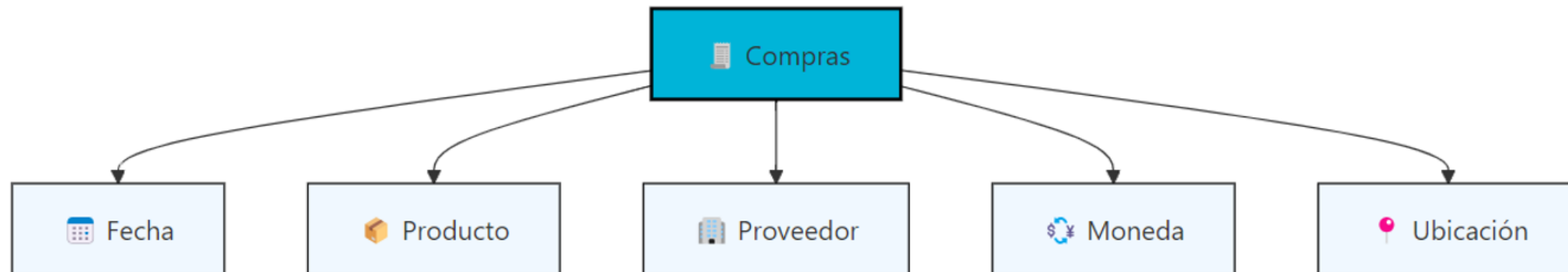
El viaje de los datos: del origen a la visualización

- Definición de la matriz de bus
- Preparación del modelo
- Flujo de los datos usando Fabric

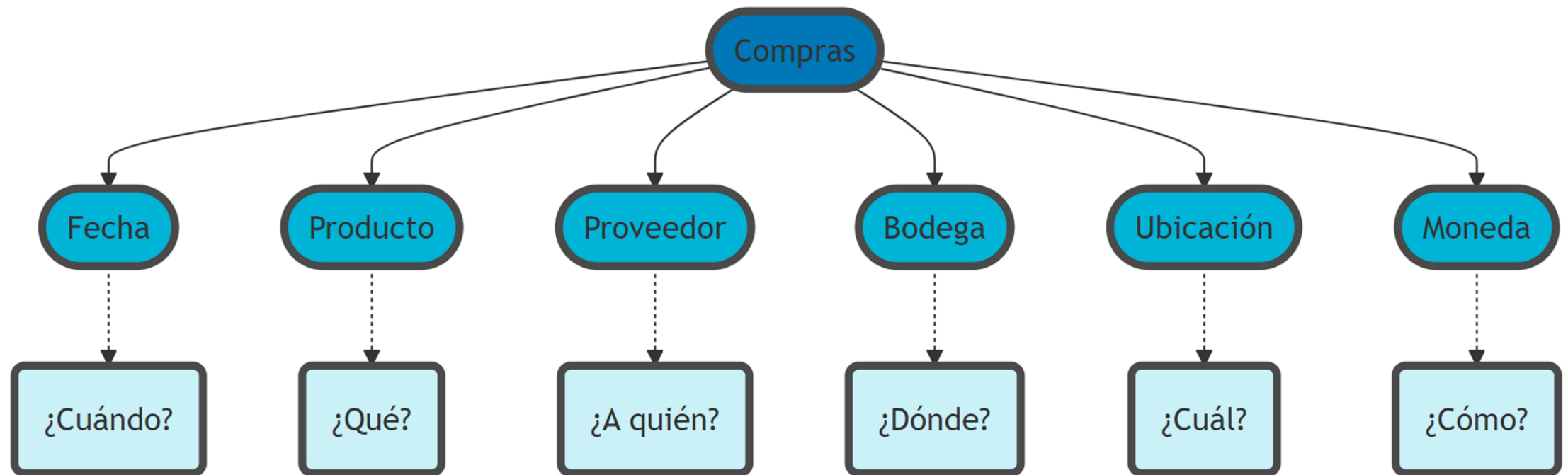


La matriz de bus

<i>Procesos de Negocio</i>	<i>Dimensiones Comunes</i>									
	Fechas	Productos	Monedas	Ubicaciones	Proveedores	Clientes	...			
Compras	✓	✓	✓	✓	✓					
Ventas	✓	✓	✓	✓		✓				
Inventarios	✓	✓		✓						



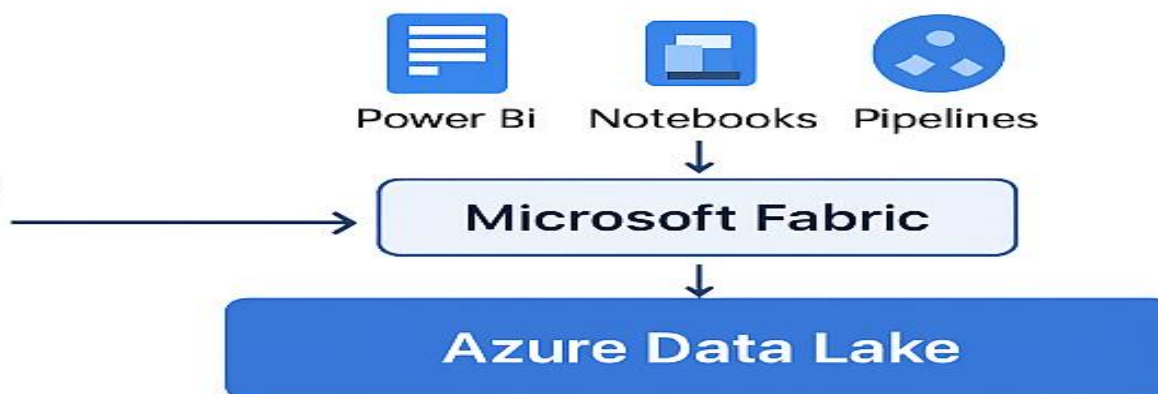
El modelo



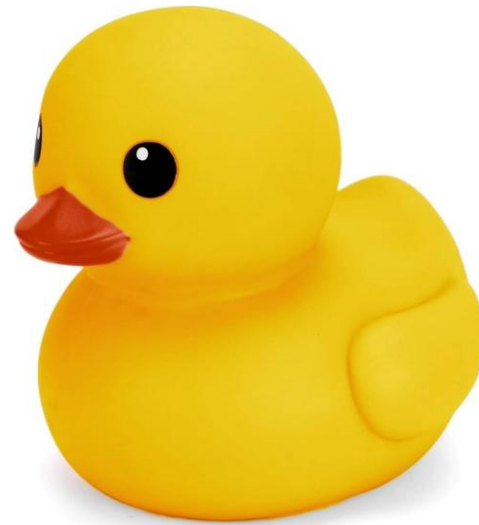
Azure Data Lake y Microsoft Fabric no compiten: se complementan.

¿Qué es?	Amacenamiento masivo de datos en la nube	Plataforma SaaS unificada para analitica
Rol principal	Lago de datos para almacenamiento estructurado y no estructurado	Orquestación, análisis y visualizacion de datos
¿Dónde se usa?	ETL, archivado, machine learning, análisis	Bi, data, science, data engineering r
Integración con Fabric	Actúa como almacenamiento base (OneLake)	Usa OneLake como capa de datos

Fabric no reemplaza a Azure Data Lake, lo potencia al agregar inteligencia, analitica y colabotación en un solo lugar



¿Qué escoger?



**DATA
WAREHOUSE**



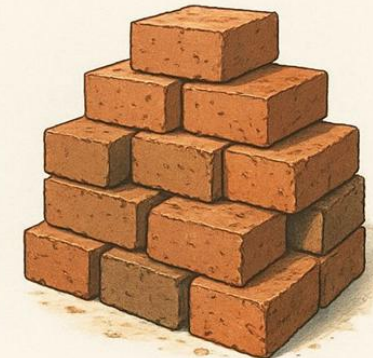
**DATA
LAKE**



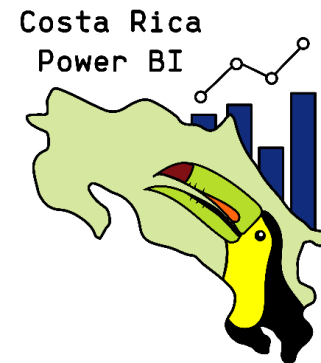
**DATA
FACTORY**



**DATA
BRICKS**



CON EL APOYO DE





Del Origen a la Visión: El Viaje de los Datos

Meibelyn Robles
Leda Araya



SJO DIGITAL 2025