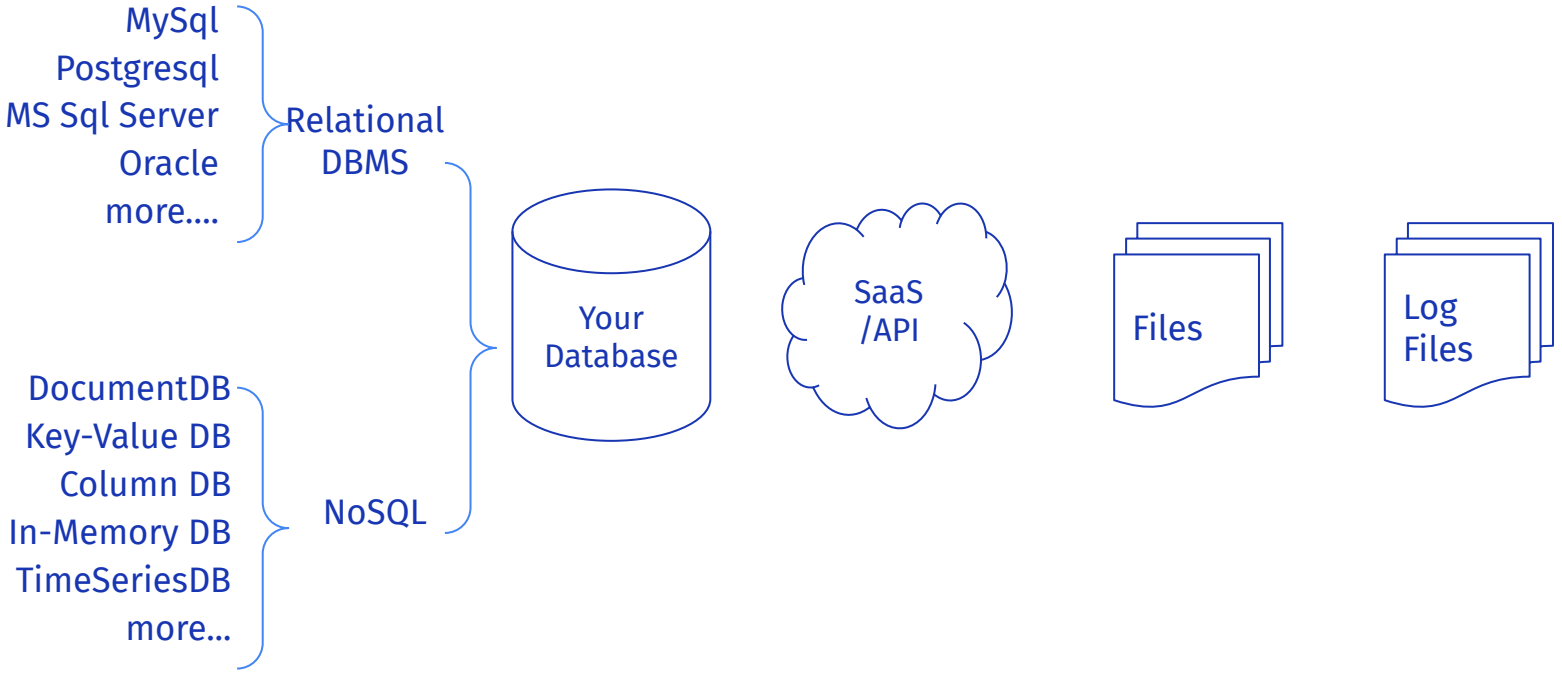
数据系统全景图

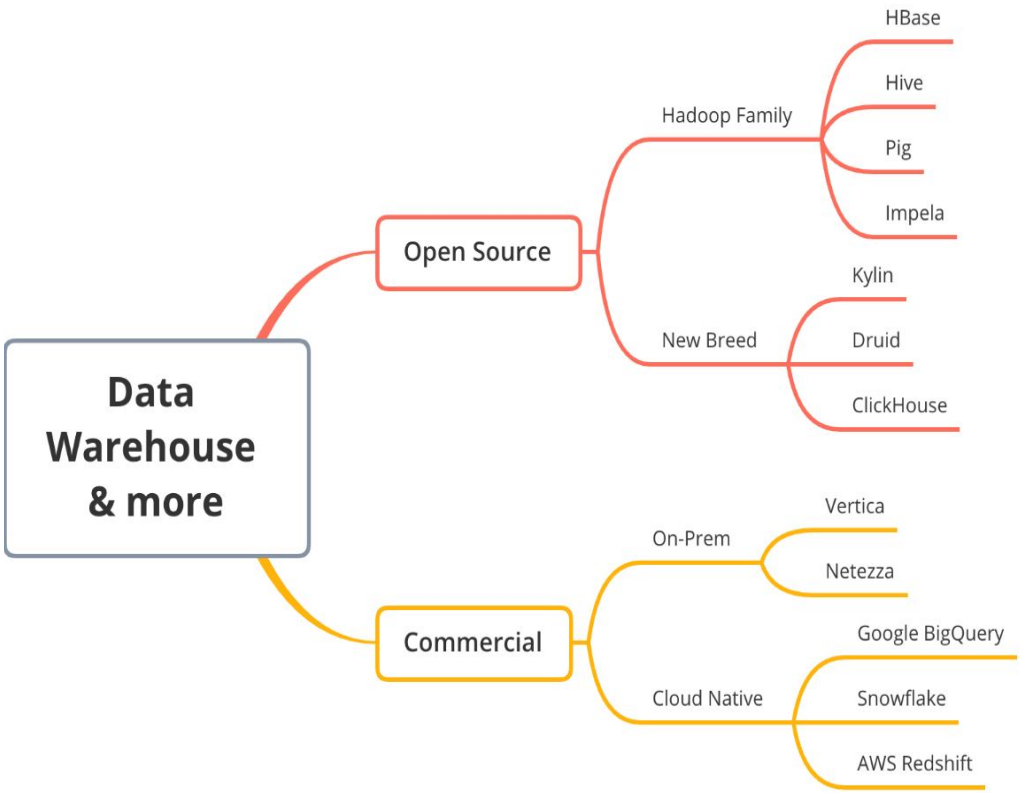
一口气说清楚数据工程的工作内容，
所用到的技术，
常见的架构和解决方案

Harry Tan

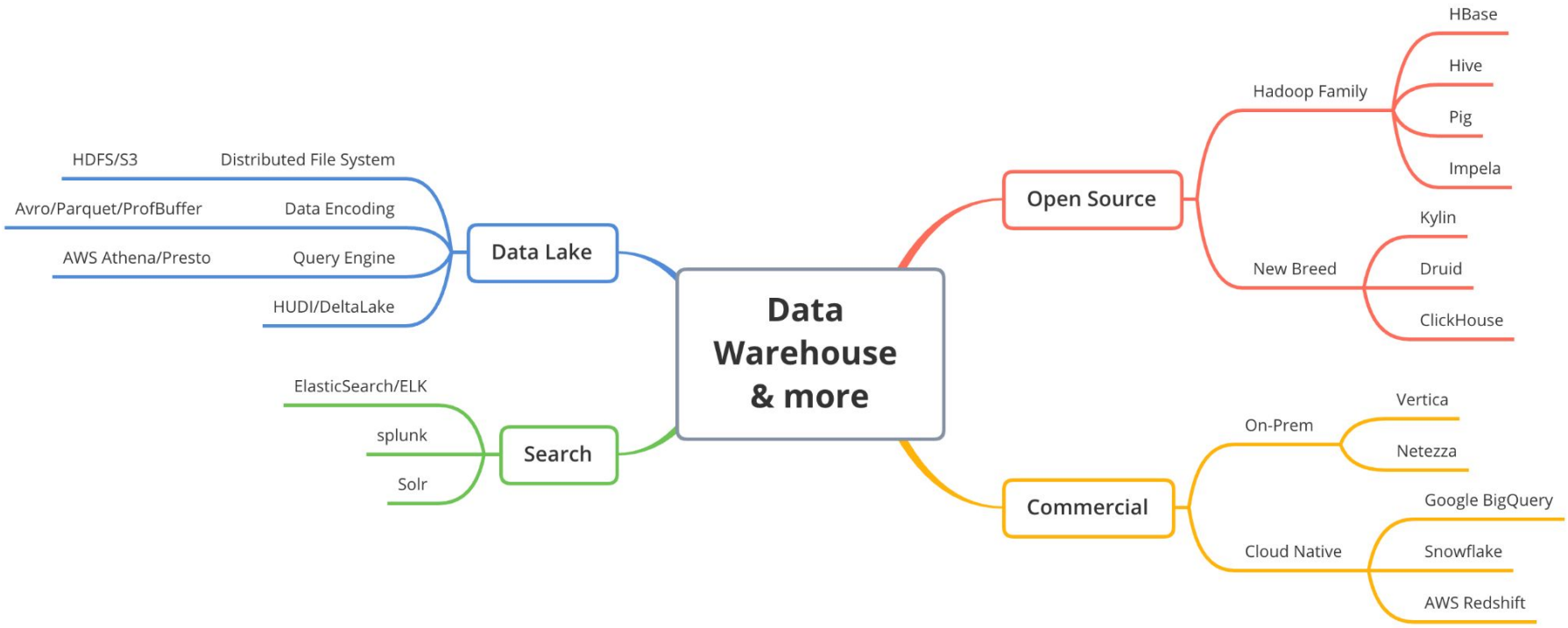
Data Source


MySql
Postgresql
MS Sql Server
Oracle
more....
Relational
DBMS
DocumentDB
Key-Value DB
Column DB
In-Memory DB
TimeSeriesDB
more...
NoSQL
Your
Database
SaaS
/API
Files
Log
Files

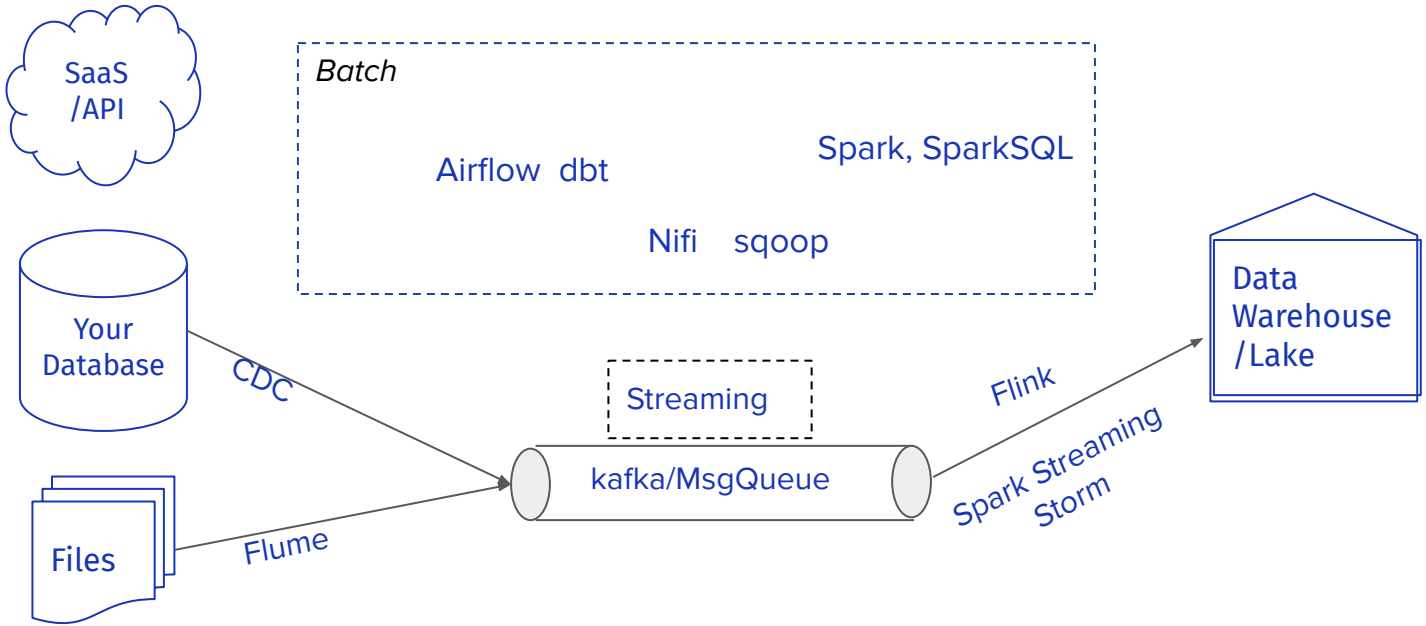
Central Data Storage



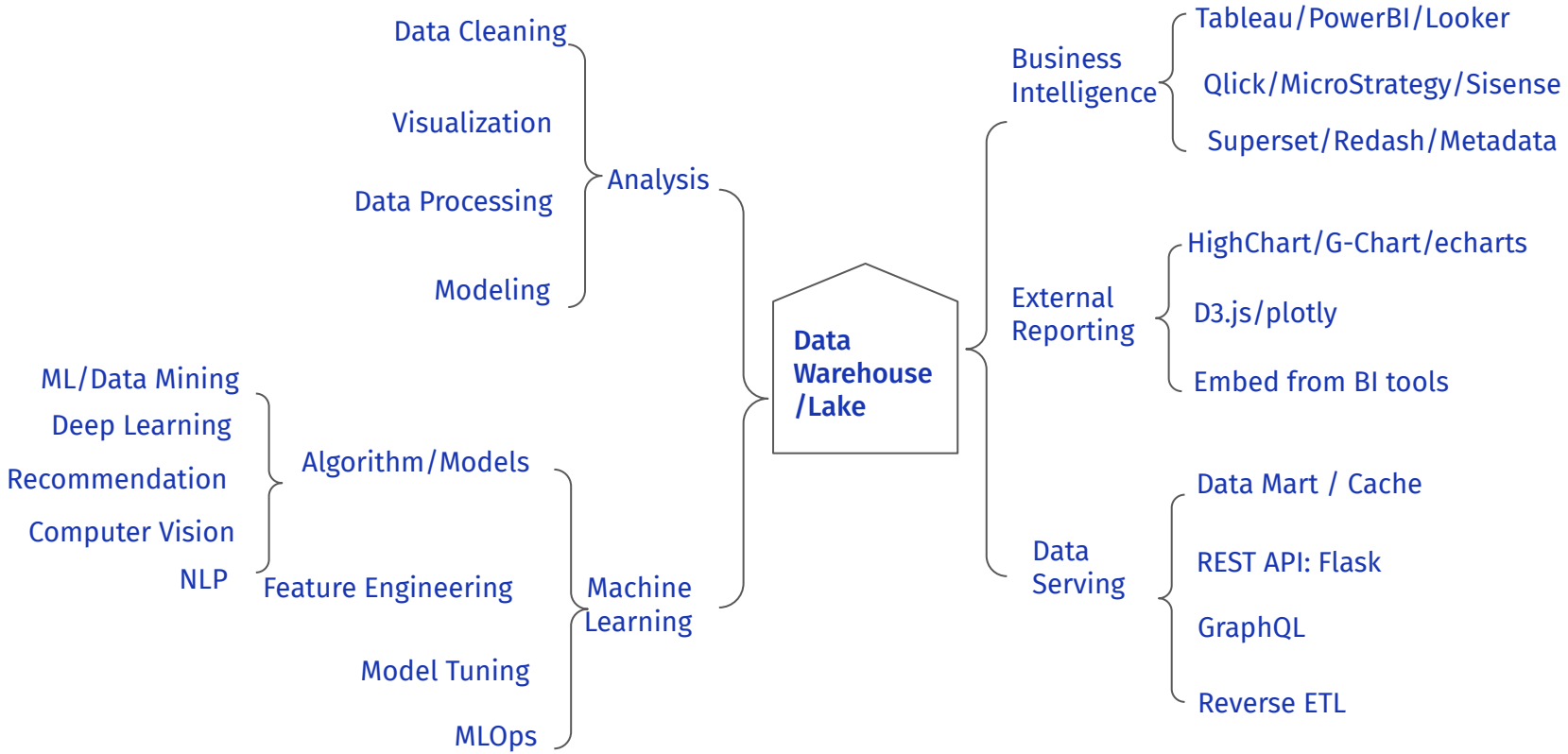
Central Data Storage



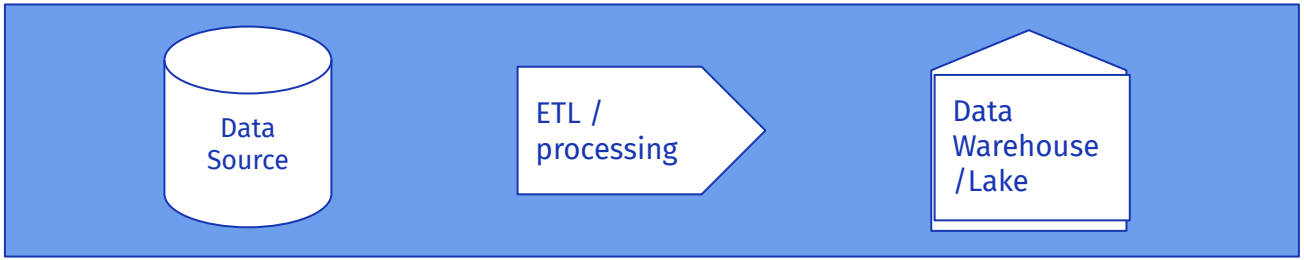
Data Pipelines

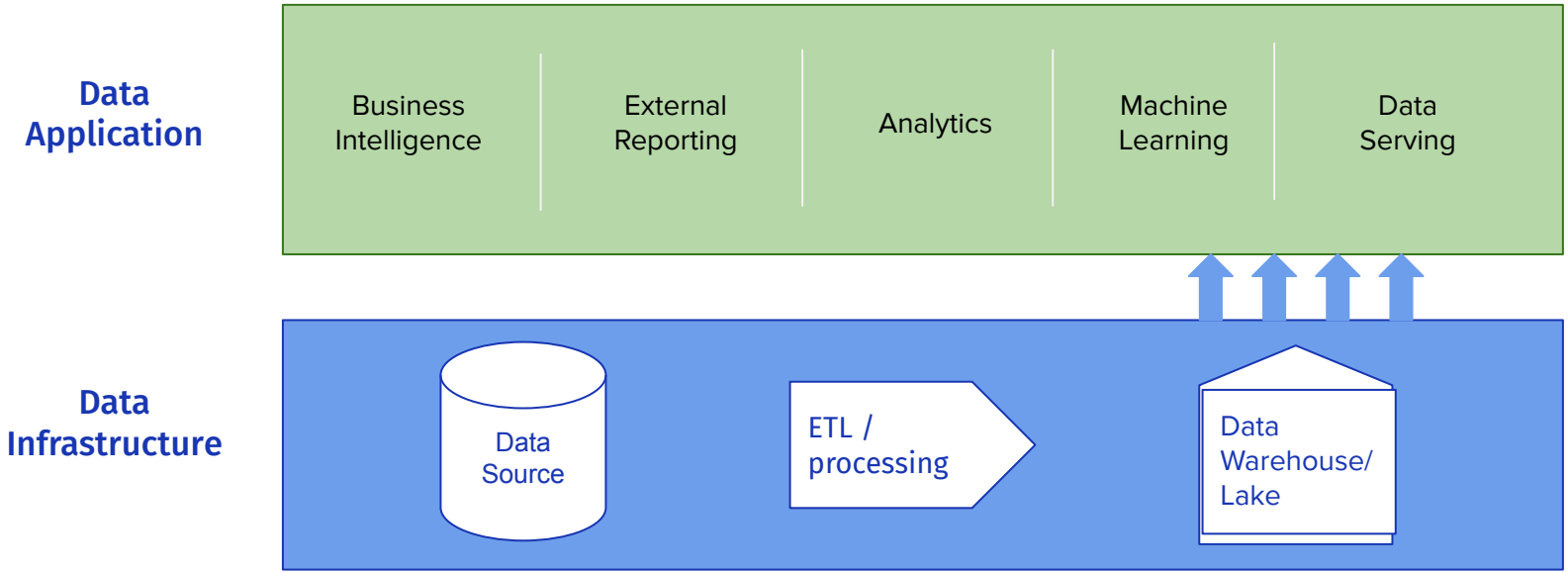


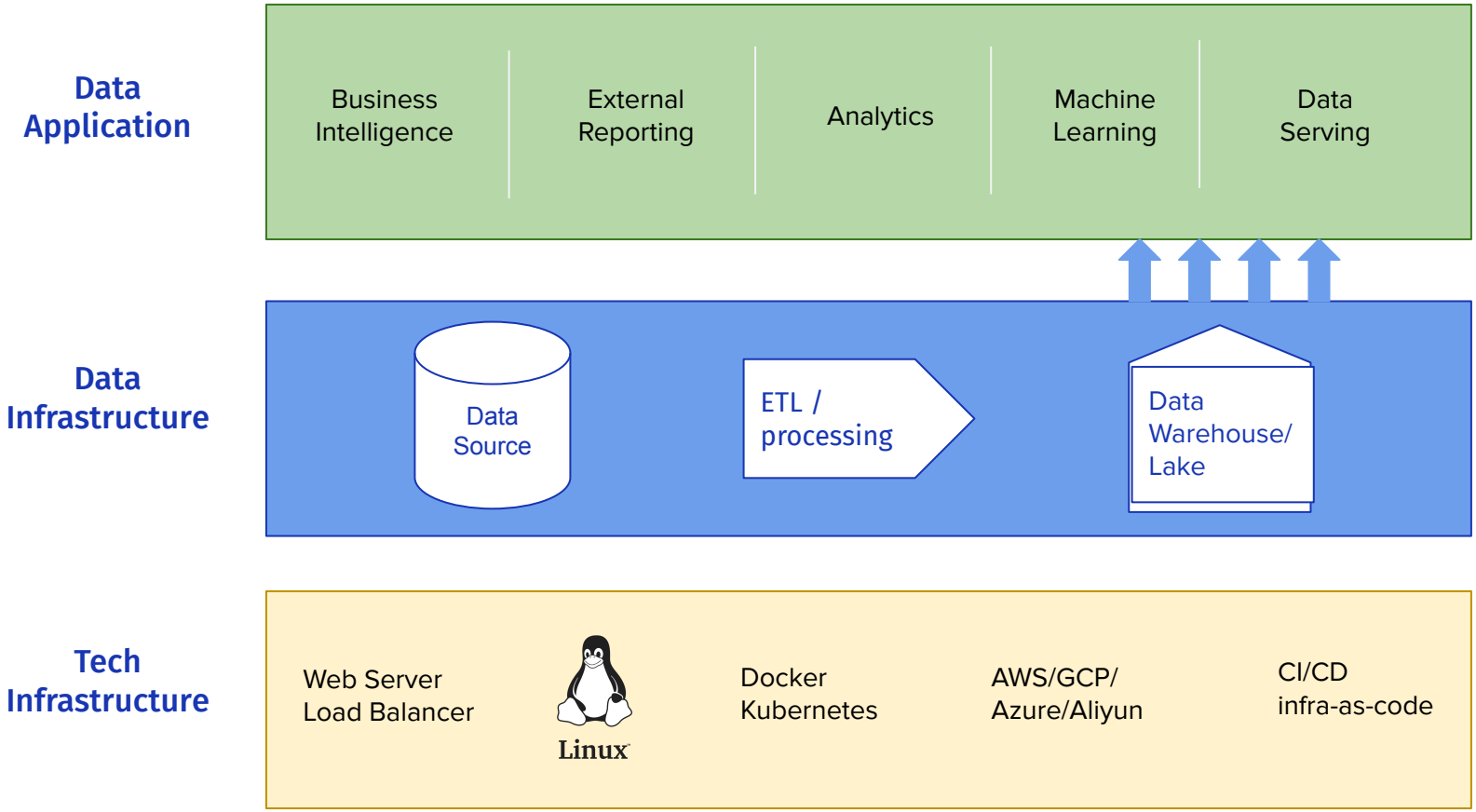
Data Applications

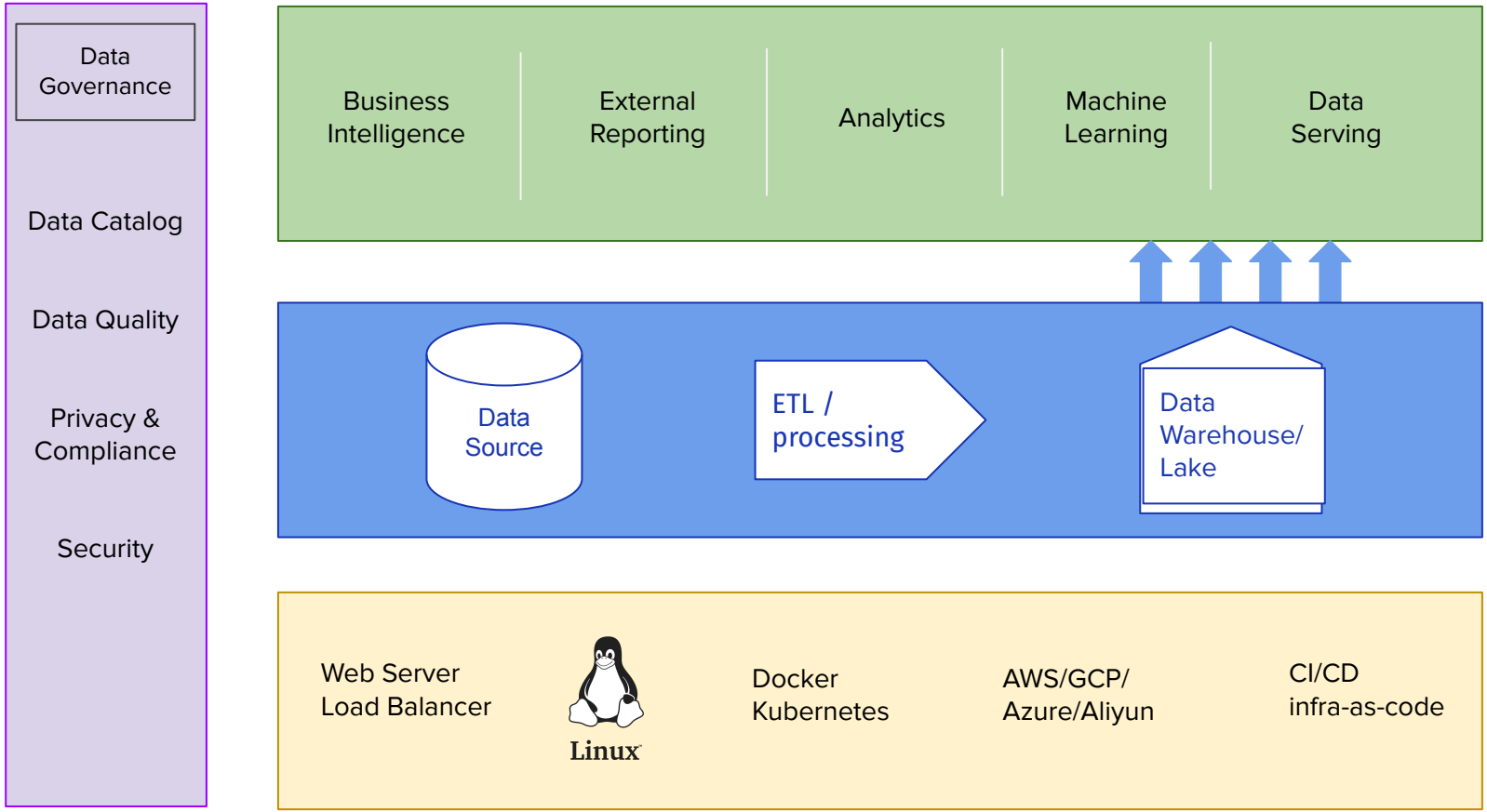


Data Infrastructure




Data Application
Business Intelligence
External Reporting
Analytics
Machine Learning
Data Serving
Data Infrastructure
Data Source
ETL / processing
Data Warehouse/ Lake


Data Application
Business Intelligence
External Reporting
Analytics
Machine Learning
Data Serving
Data Infrastructure
Data Source
ETL / processing
Data Warehouse/ Lake
Tech Infrastructure
Web Server
Load Balancer
Linux
Docker
Kubernetes
AWS/GCP/
Azure/Aliyun
CI/CD
infra-as-code


Data Governance
Data Catalog
Data Quality
Privacy & Compliance
Security
Business Intelligence
External Reporting
Analytics
Machine Learning
Data Serving
Data Source
ETL / processing
Data Warehouse/ Lake
Web Server
Load Balancer
Linux
Docker
Kubernetes
AWS/GCP/
Azure/Aliyun
CI/CD
infra-as-code