BAYER'S ALYCE: ADVANCING DRUG DEVELOPMENT WITH DATA INTELLIGENCE

Dirk Helwig

CDS & A
Clinical Data Sciences & Analytics

June 2024

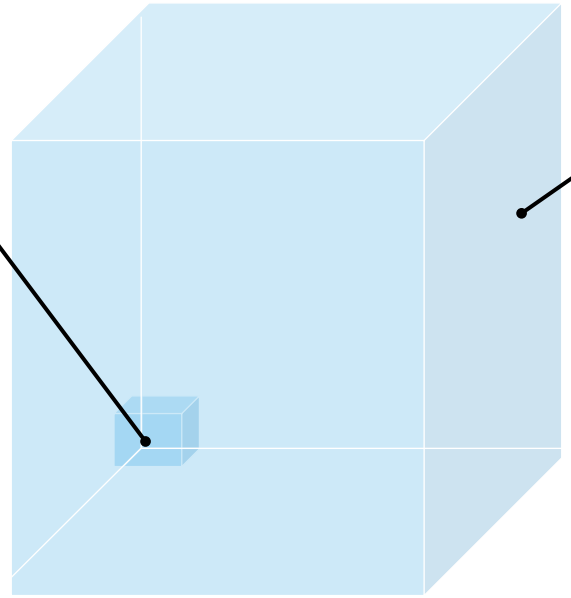
DATA+AI SUMMIT BY databricks

Decentralized Clinical Trials

ALYCE
ADVANCED ANALYTICS PLATFORM for the CLINICAL ENVIRONMENT

Data collected at on-site visits

- // Demographics
- // Weight, Height
- // Vital Signs
- // Single 12 Lead ECGs
- // Adverse Events, Medical History
- // …

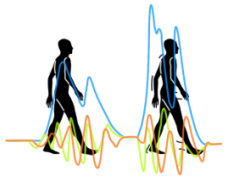


Volume, Speed and Complexity of generated data will increase

Data from Digital Health Technologies collected at optimal time points e.g.:

Actigraphy devices

- // Physical Activity
- // Energy Expenditure
- // Sleep Activity
- // Body Posture

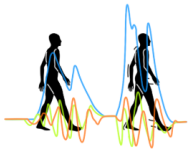


ECG Patches

- // Heart Rate
- // Respiratory Rate
- // R-to-R Interval
- // Fluid Status

Ingesting / Wrangling / Analysing

Data ingestion (1)



Activity data

Sleep data

ECG data

eCOA

other data sources

(2) Data orchestration, Reports / Dashboards & Advanced Analytics

// Ingestion
// Quality check
// Merge, Wrangling…
// Visualizations
// Notifications
// End-2-End Pipelining
// Advanced Analytics

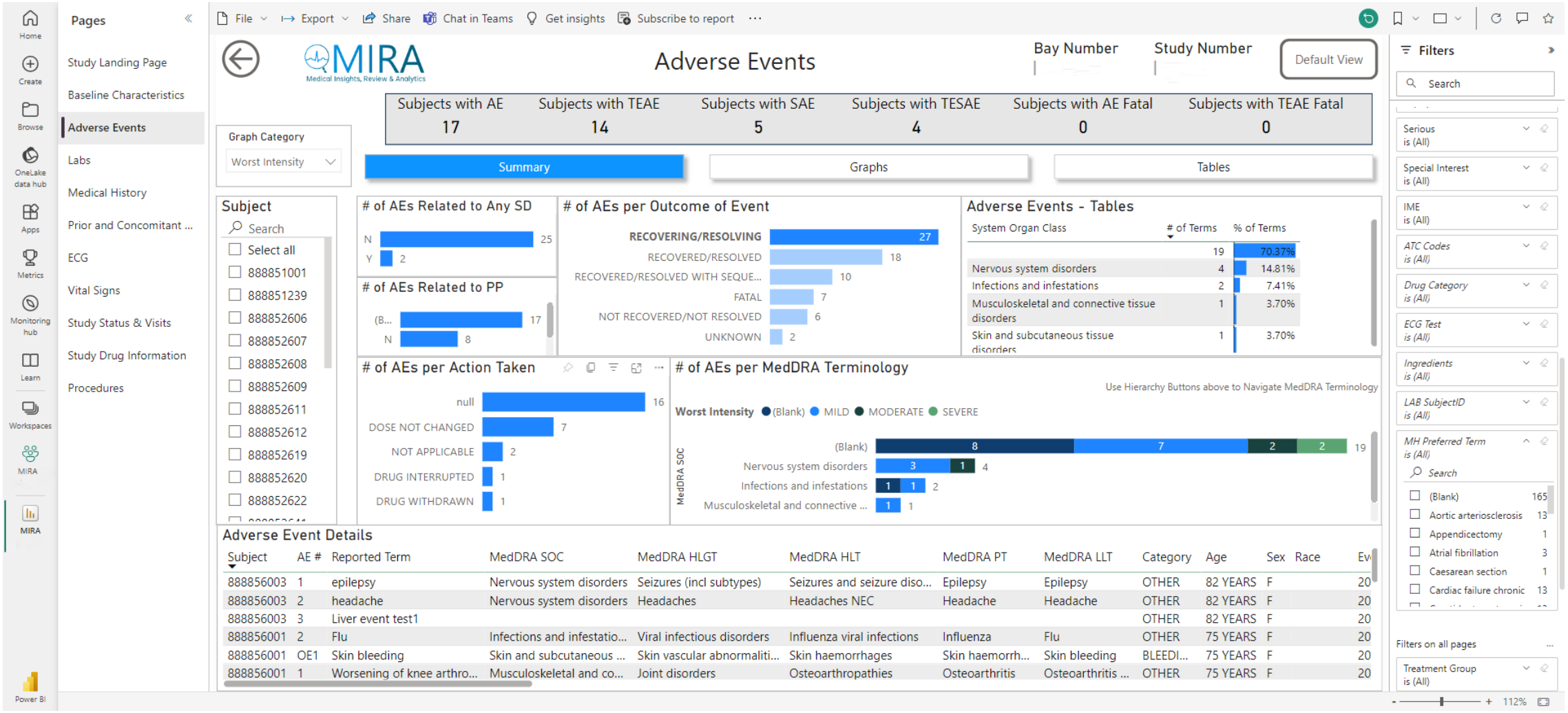
ALYCE ADVANCED ANALYTICS PLATFORM for the CLINICAL ENVIRONMENT

(3) ***Push of Processed Data***

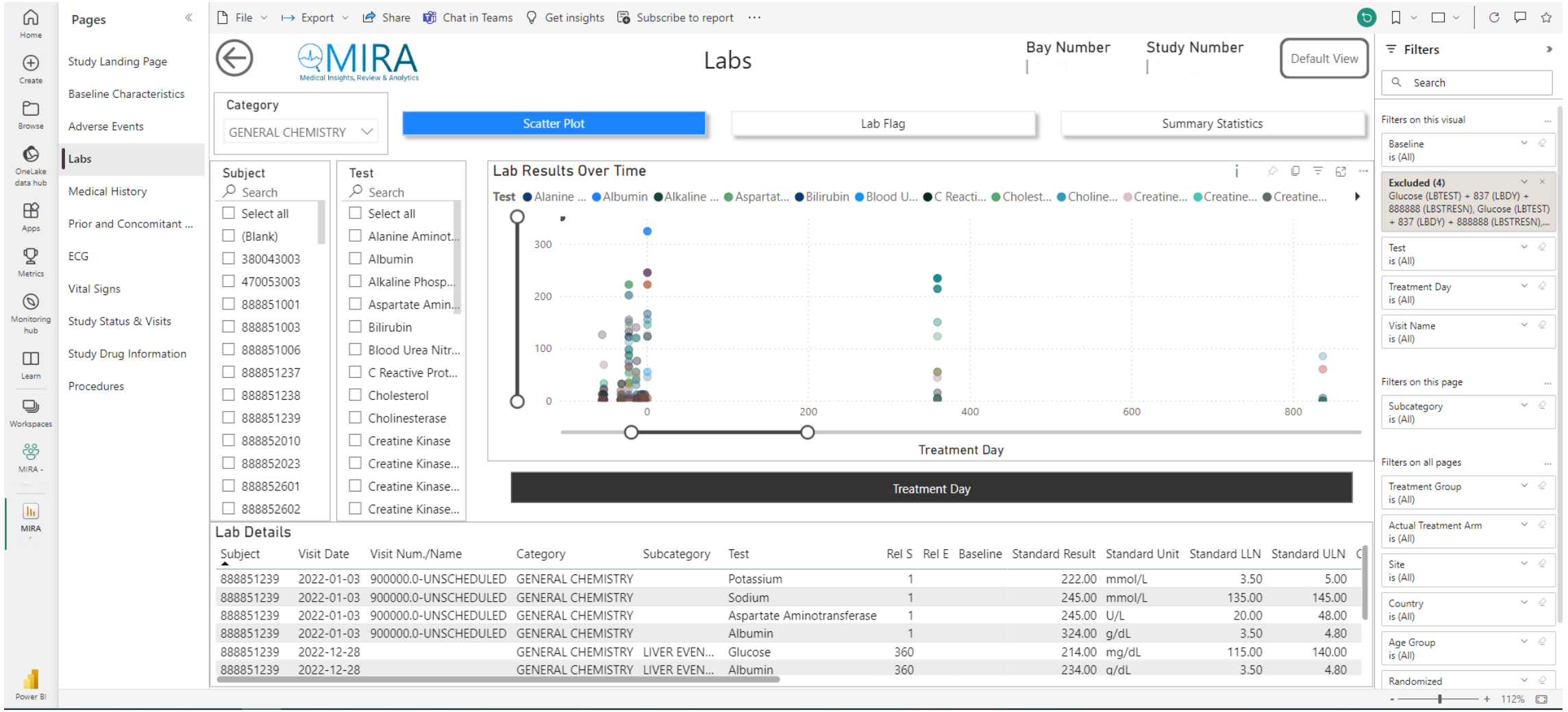
Operational database

(Clinical Data Warehouse)

Therapeutic Area Overarching (TAO) dashboards (AE)

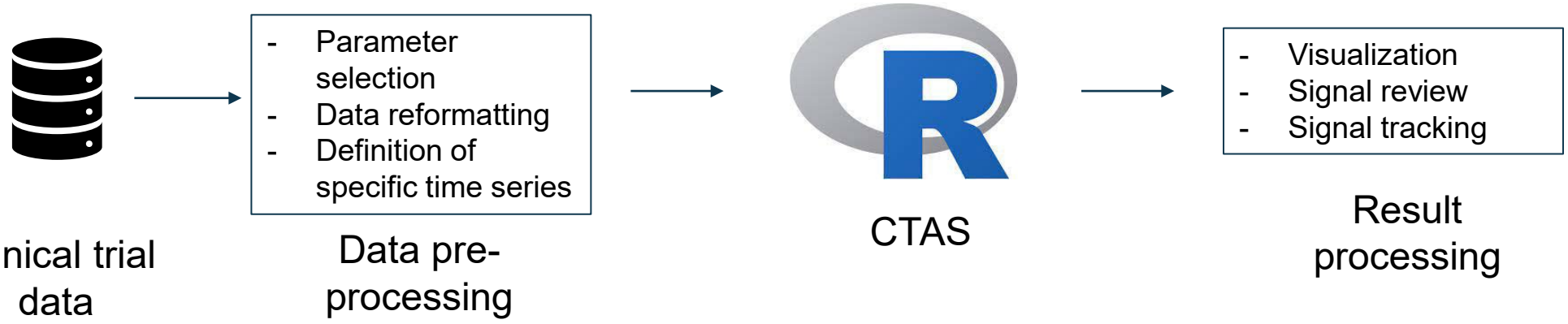


Therapeutic Area Overarching (TAO) dashboards (Labs)



Clinical Trial Anomaly Spotter (CTAS)

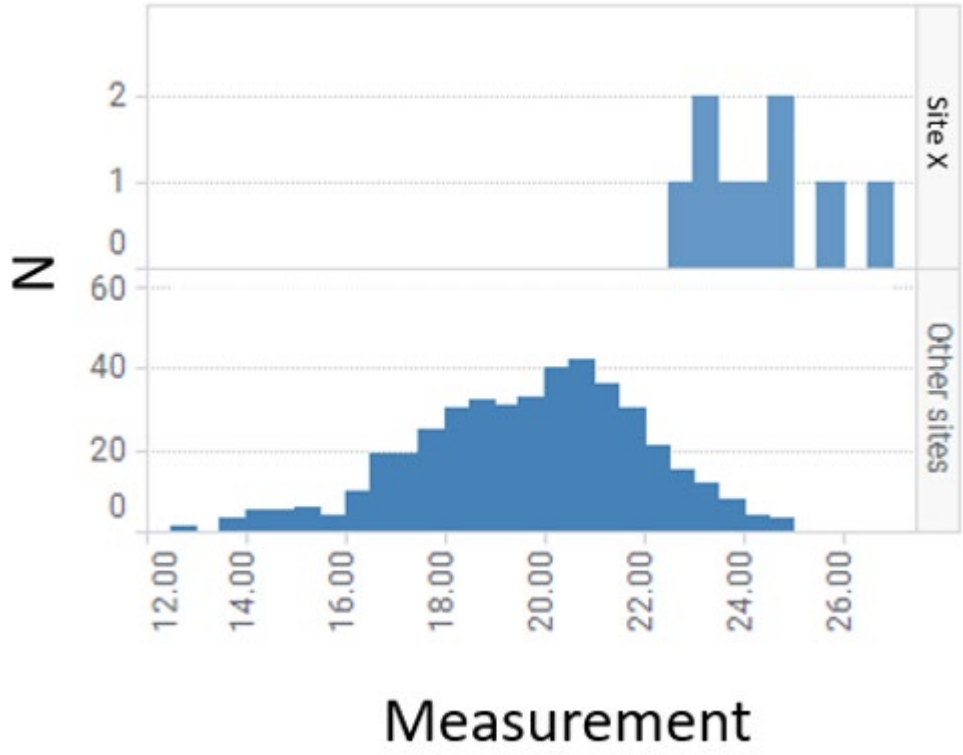
- CTAS (Clinical Trial Anomaly Spotter) is a CSM (Central Statistical Monitoring) tool focused on identifying anomalous time series of continuous variables.
- “Unsupervised” tool which does not use pre-specified KRIs. The identified anomalies can be such that were not anticipated during study design.
- Originally developed by Bayer where it is available for all company studies.
  - The Bayer-internal tool was presented in PHUSE EU Connect 23. Please see refs. [1] and [2] for details.
- Code base shared with IMPALA for co-development in July 2023.



[1] https://phuse.s3.eu-central-1.amazonaws.com/Archive/2022/Connect/EU/Belfast/PRE_AR04.pdf
[2] https://phuse.s3.eu-central-1.amazonaws.com/Archive/2022/Connect/EU/Belfast/PAP_AR04.pdf

Examples of anomalies identified with CTAS

Site with unusually large laboratory assay results (Bayer)



Potential reasons: bias in subject selection, in-correct assay calibration (local labs), errors in sample handling (central labs)

IMPALA consortium

- Industry consortium IMPALA has been formed to develop tools, provide training and create best practices for quality in clinical trials.
- IMPALA aims to engage with Health Authorities inspectors on defining guiding principles for the use of advanced analytics to complement, enhance and accelerate current QA practices.
- Another important activity is co-development of open-source AI/ML tools for QA.
- For more information on the IMPALA consortium,
  please visit https://impala-consortium.org/

CTAS is open source

After few months of co-development, IMPALA published CTAS as an open-source R package in December 2023.

The source code is available on the public IMPALA Github repository under the MIT License:
https://github.com/IMPALA-Consortium/ctas

Everyone is invited to test and use the package and report bugs and/or suggest new features.

Azure Databricks, OpenAI & Langchain for clinical GenAI experiments

Microsoft Azure | databricks

Search data, notebooks, recents, and more... CTRL + P

New
Workspace
Recents
Catalog
Workflows
Compute

SQL
SQL Editor
Queries
Dashboards
Alerts
Query History
SQL Warehouses

Data Engineering
Job Runs
Data Ingestion
Delta Live Tables

Machine Learning
Playground
Experiments
Features
Models
Serving

Marketplace
Partner Connect

Collapse menu

LLMs for clinical (statistics) examples Python

File Edit View Run Help Last edit was now New cell UI: ON

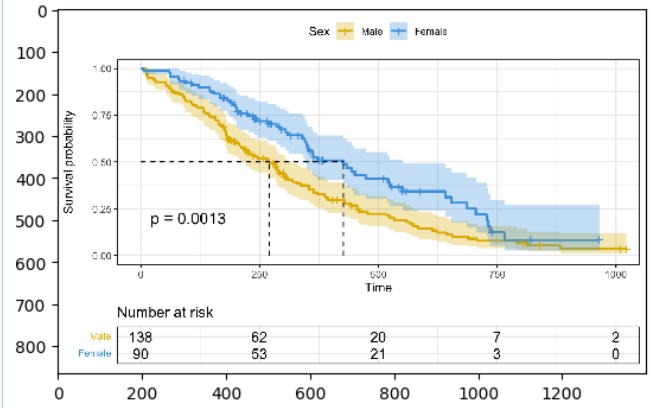
Run all | Terminated | Schedule | Share

Markdown

Use of GPT4-vision to interprete clinical figures

2/22/2024 (1s) 26: Examples KM curve from survminer

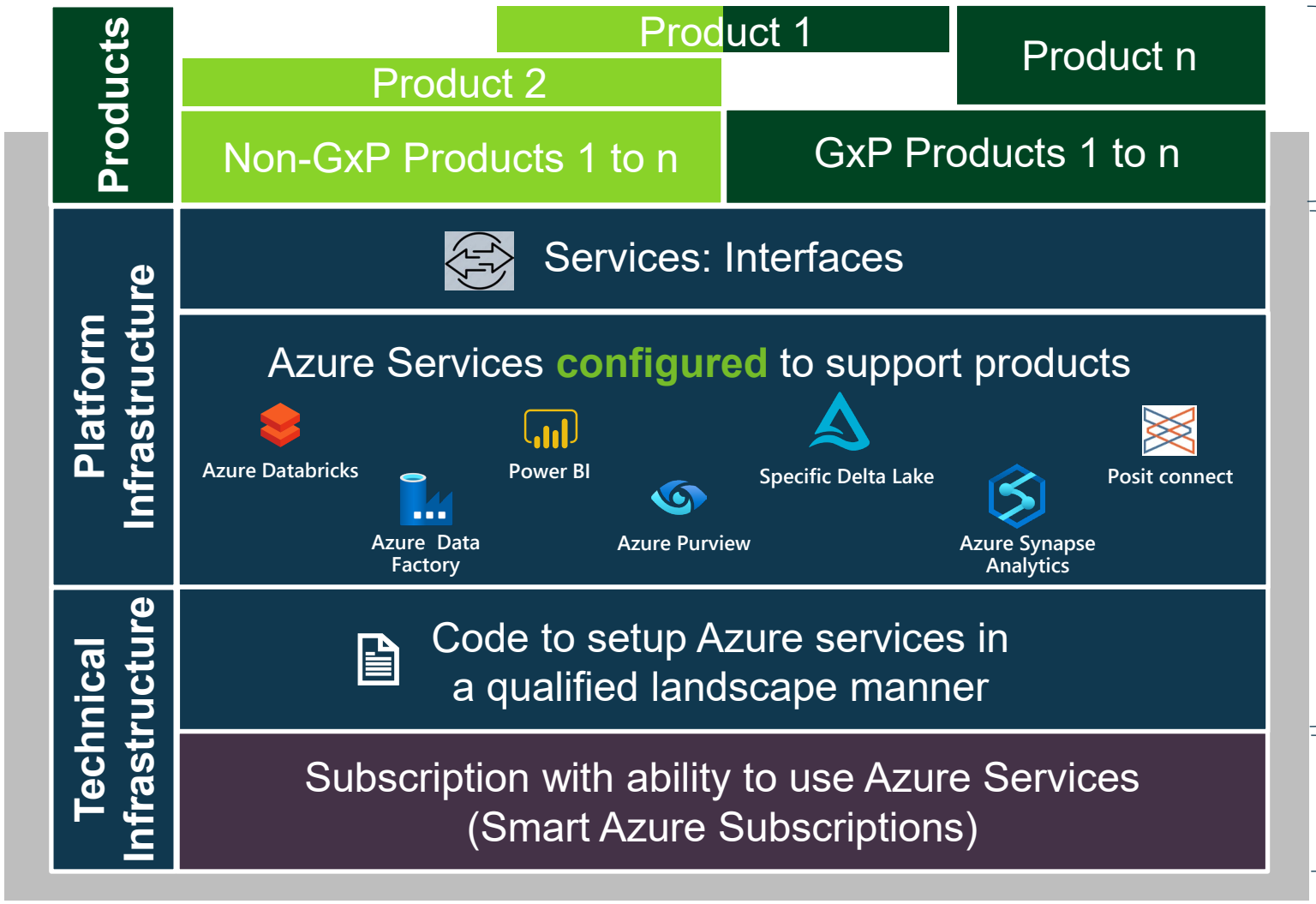
```
1  import base64
2  import io
3  from matplotlib import pyplot as plt
4  import matplotlib.image as mpimg
5  import requests
6  b64 = base64.b64encode(requests.get("https://rpkgs.datanovia.com/survminer/reference/ggsurvplot-2.png").content)
7  plt.imshow(mpimg.imread(io.BytesIO(base64.b64decode(b64))), interpolation='nearest')
8  plt.show()
```



2/22/2024 (10s) 27: Interprete a Kaplan Meier / survival curve with full response

```
1  chat_completion = vision_client.chat.completions.create(
```

ALYCE Platform GxP Blueprint



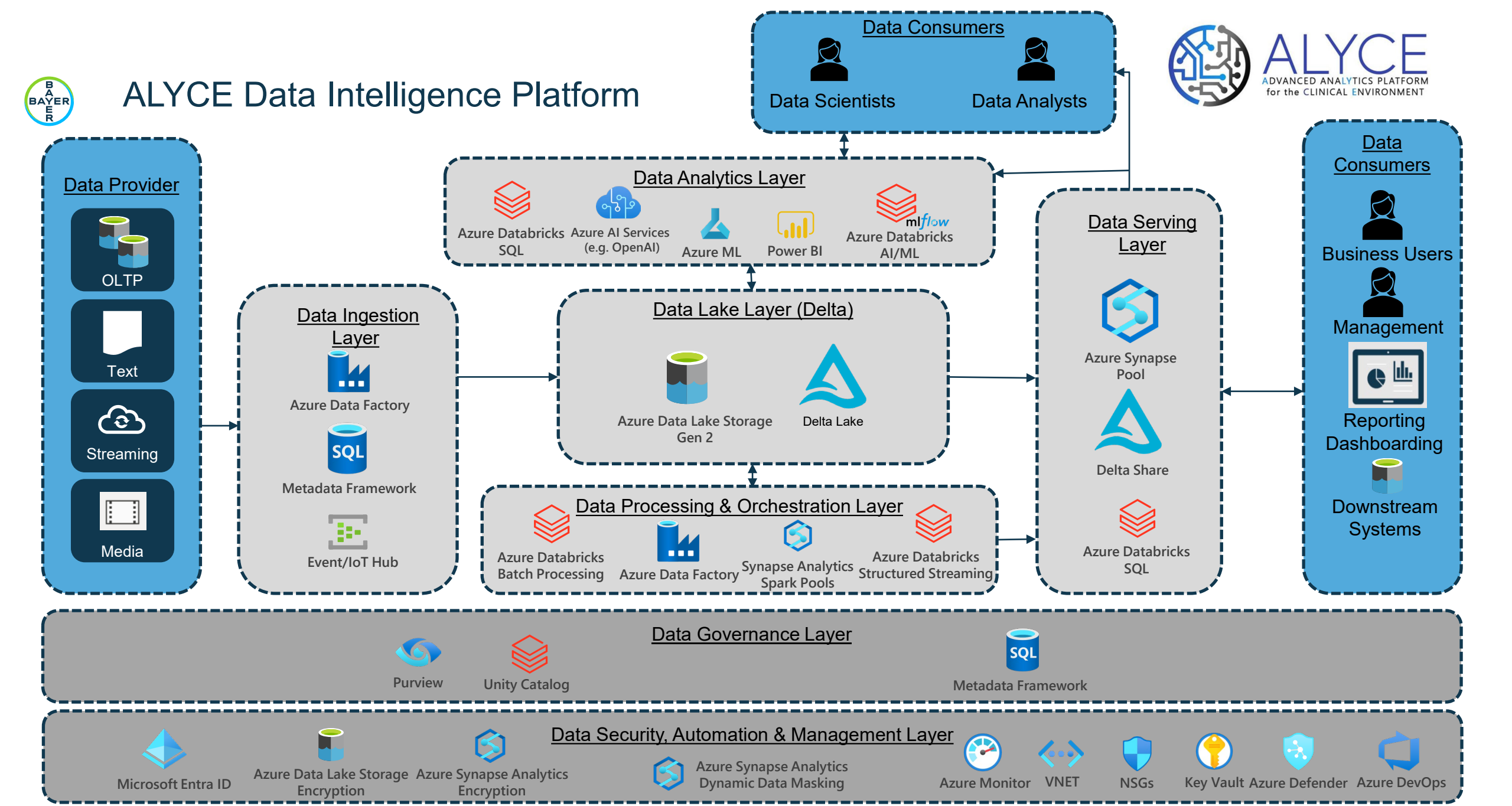
ALYCE Product Validation

ALYCE Platform Qualification (IaC)

Qualified Bayer Subscriptions

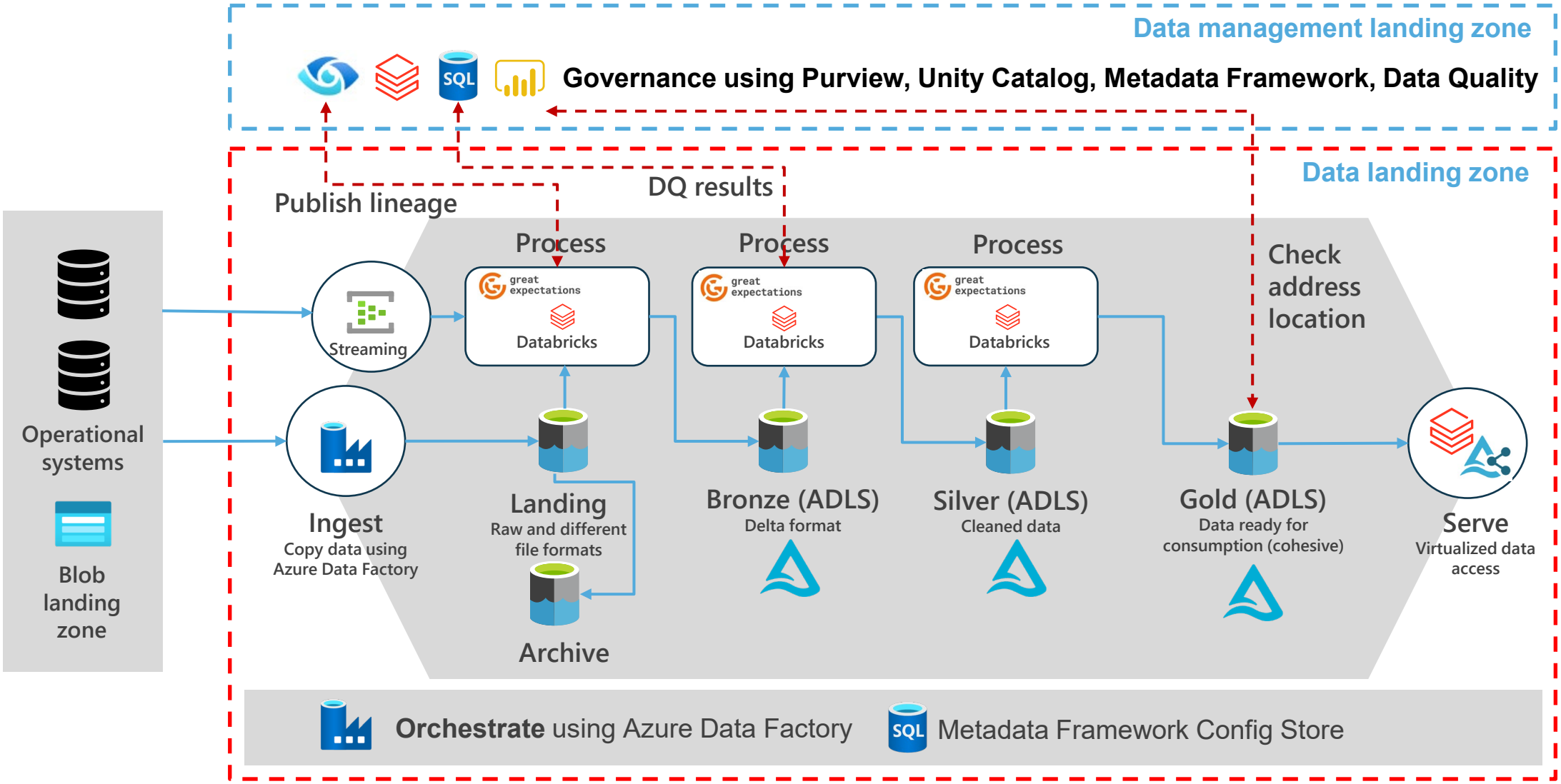
Qualified Azure DevOps & Modern Requirements

ALYCE Data Intelligence Platform



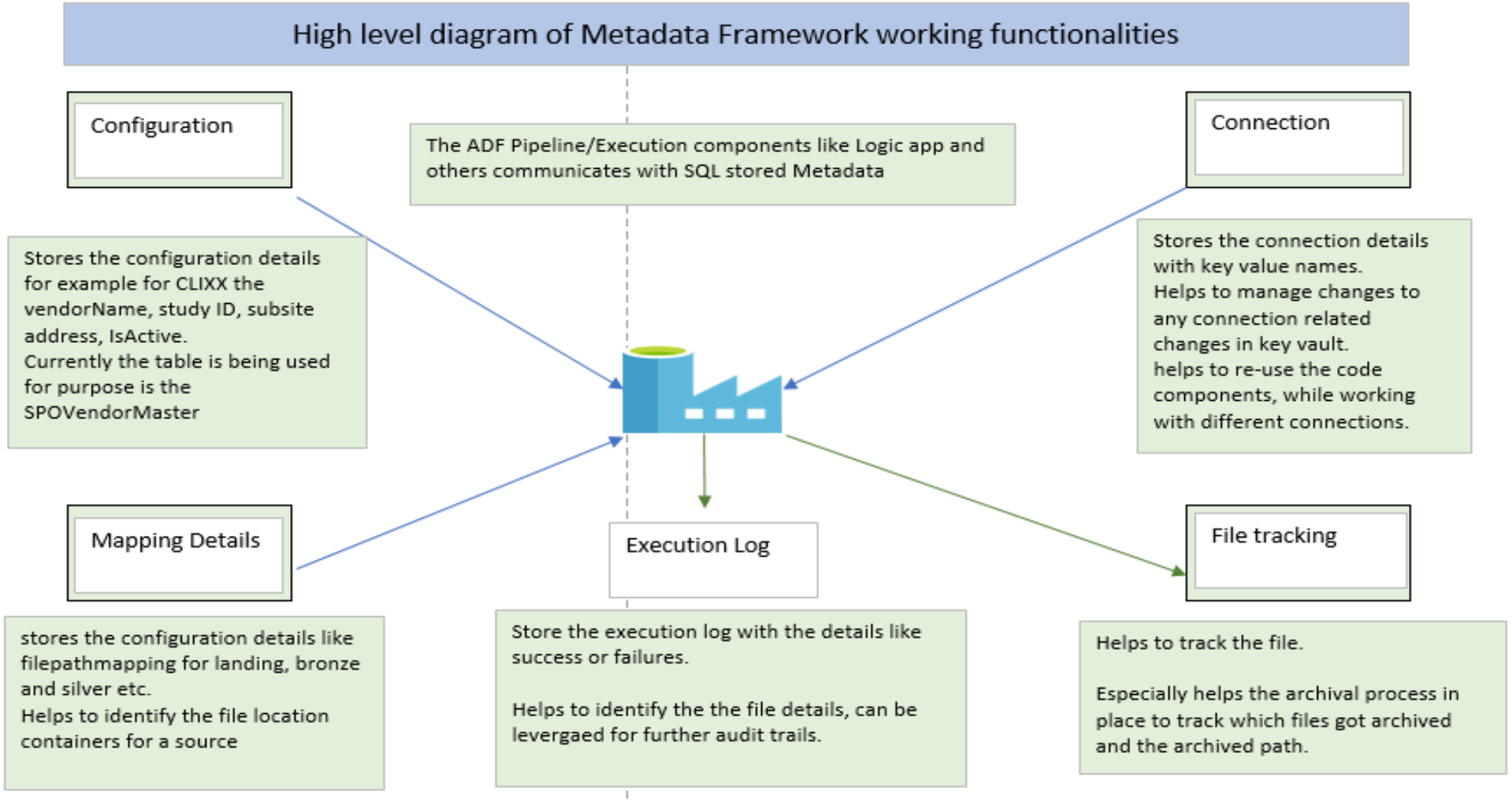
ALYCE Data Product Blueprint

Each domain get its own data product architecture for building data products at scale

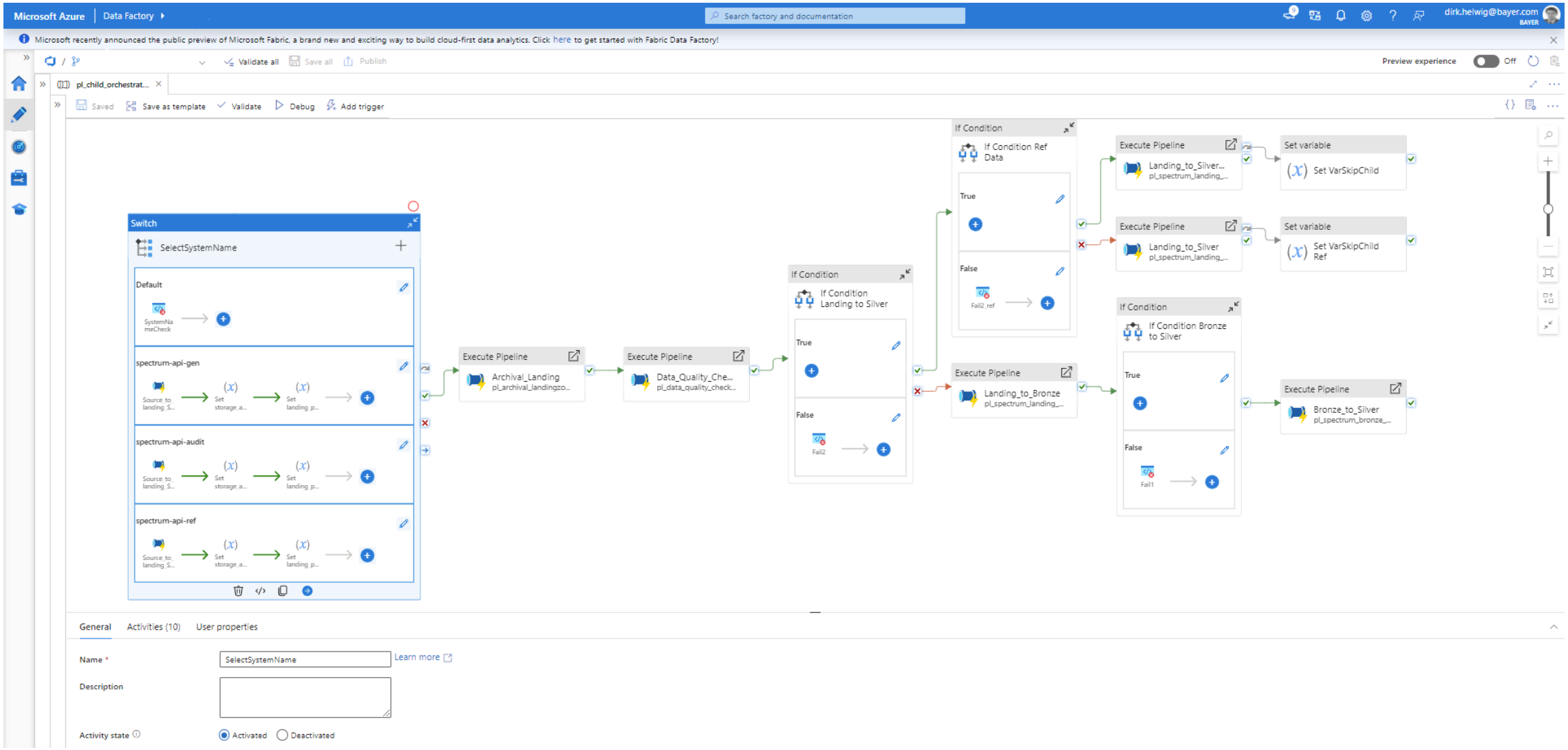


Metadata Framework Config Store

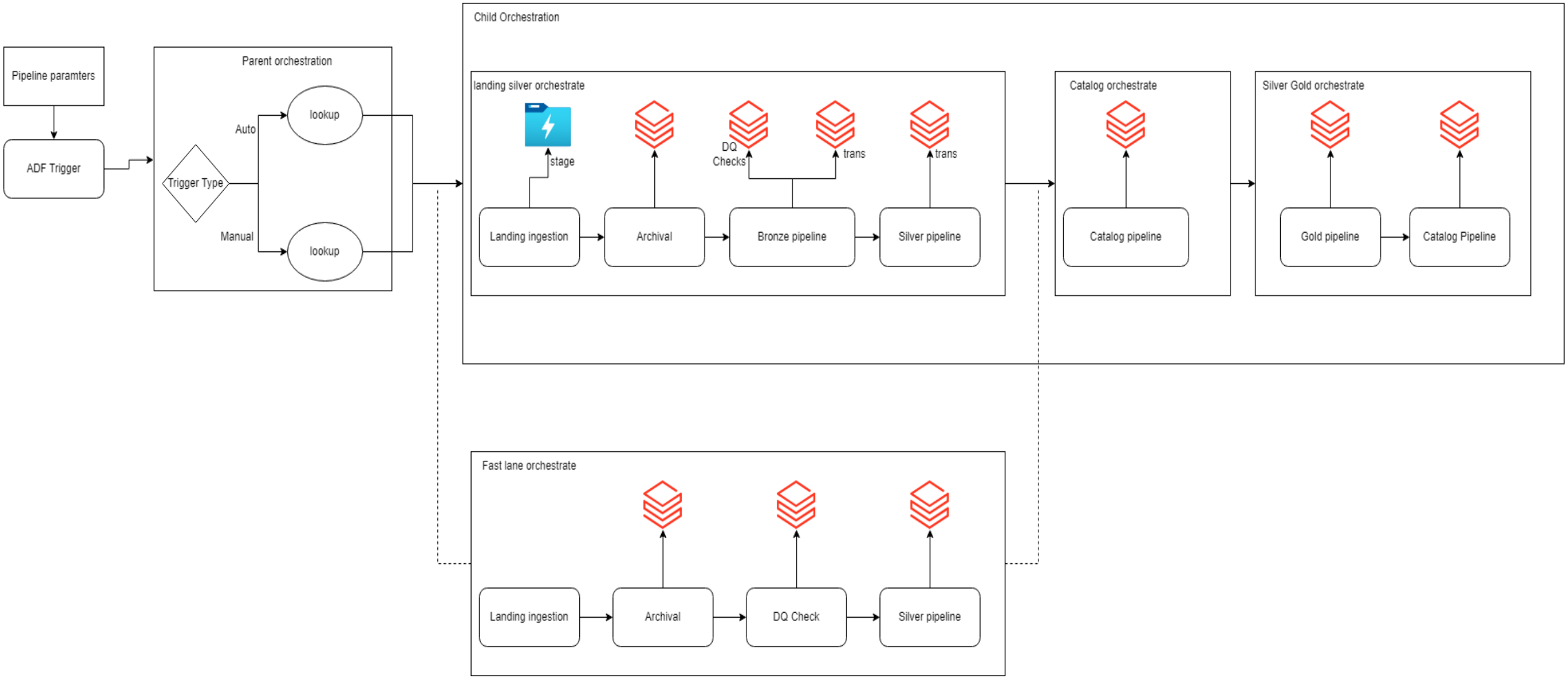
Implementation design of Metadata Framework



Azure ADF Generalized Pipeline

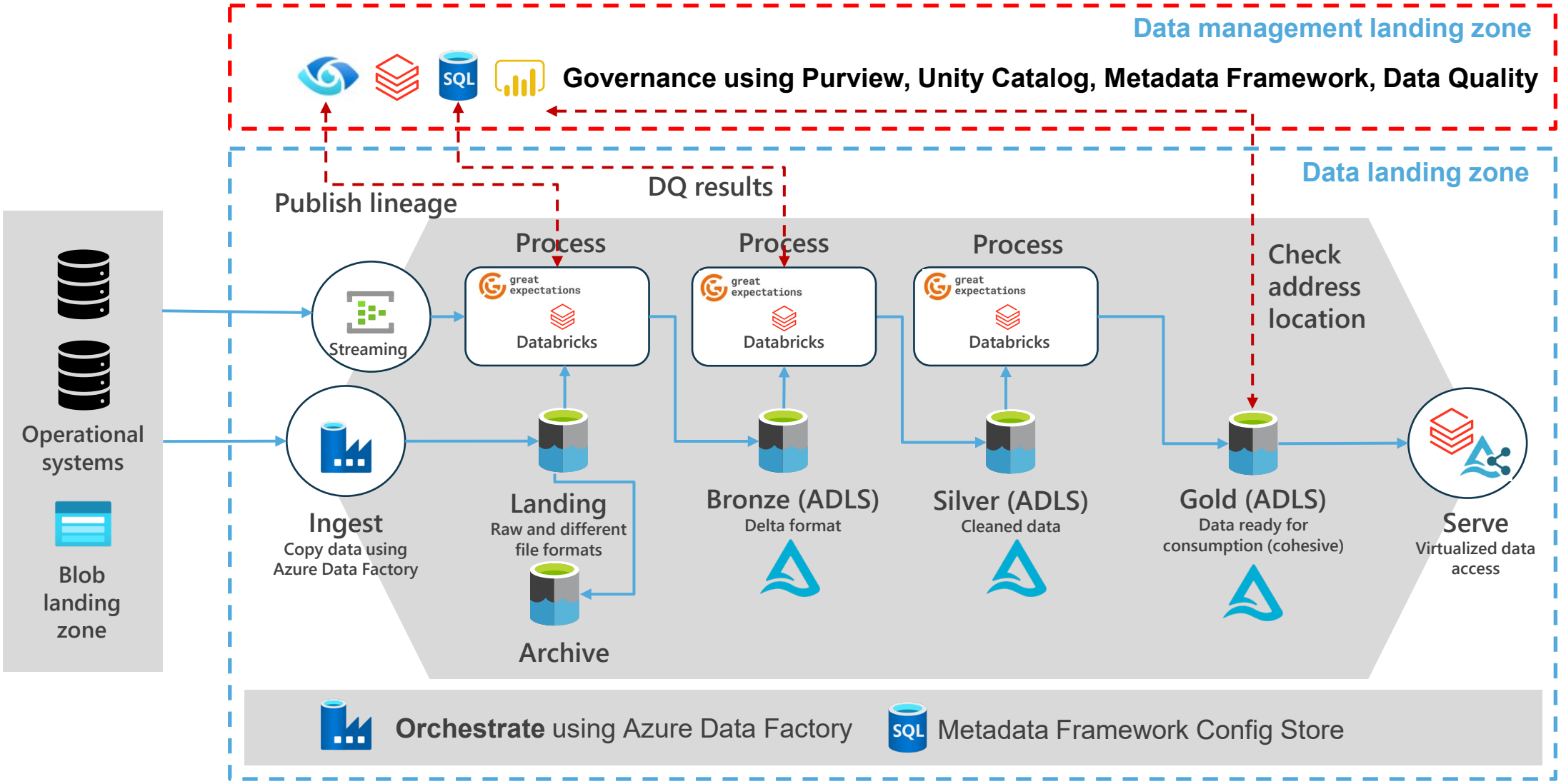


ALYCE Orchestration Blueprint

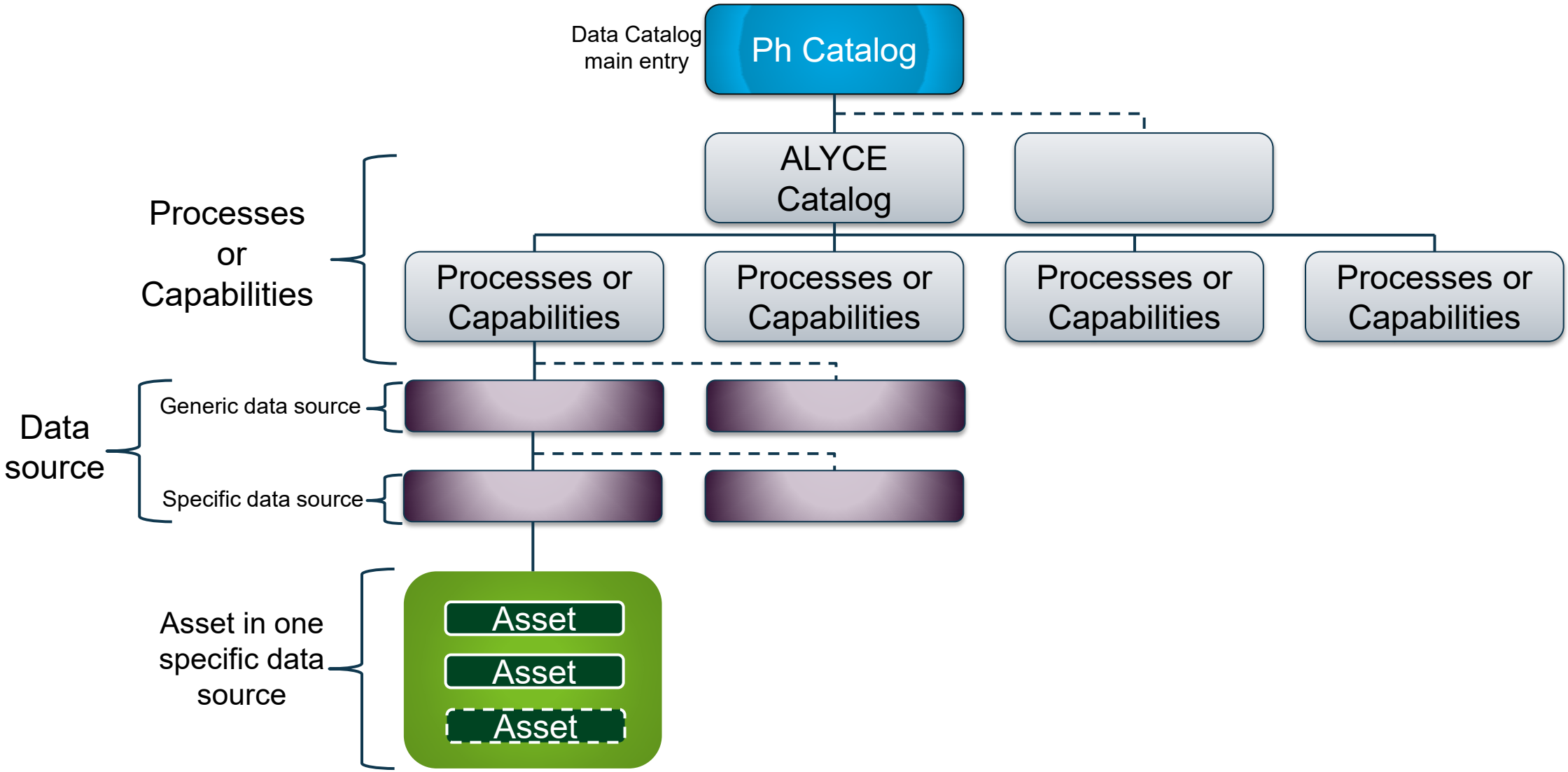


ALYCE Data Product Blueprint

Each domain get its own data product architecture for building data products at scale

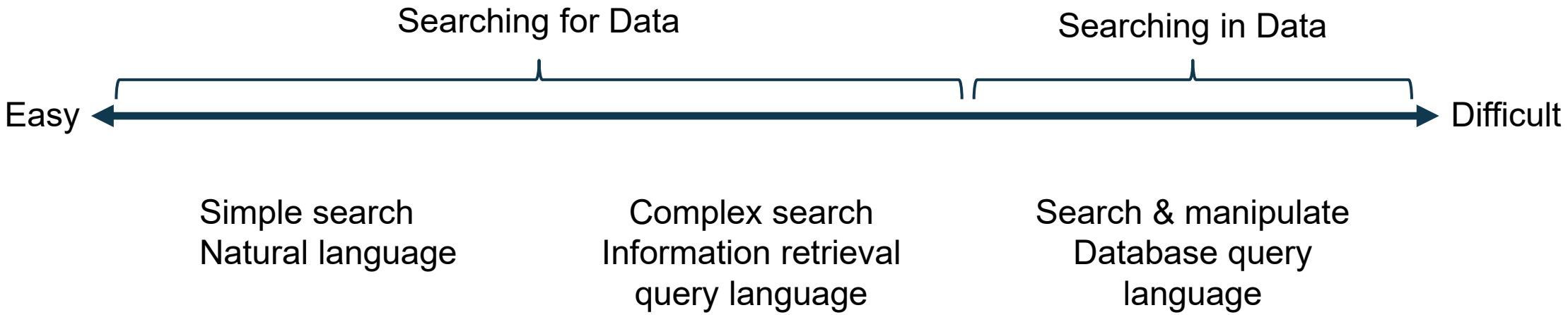


ALYCE Data Catalog Blueprint

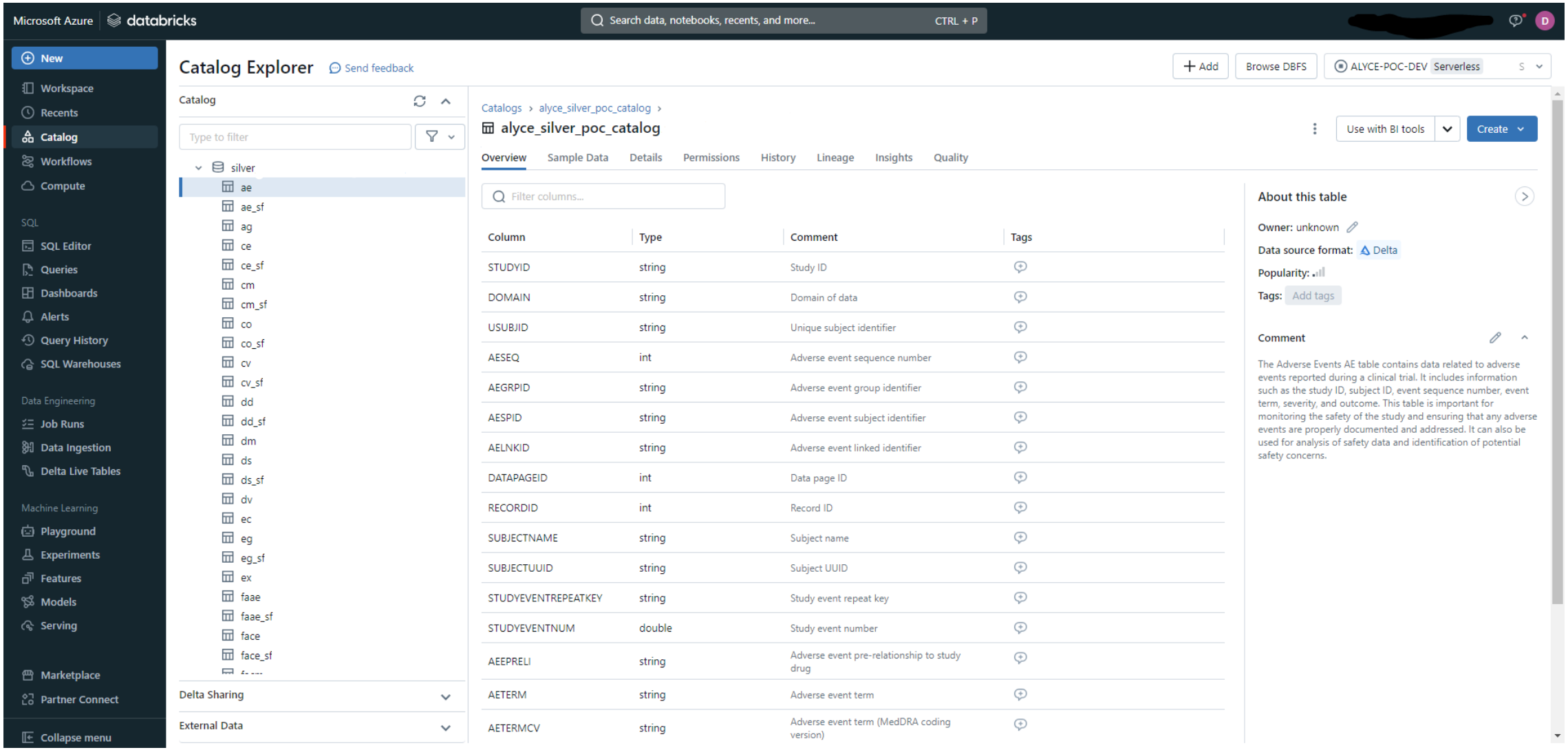


ALYCE Data Catalog Search

Spectrum of Search

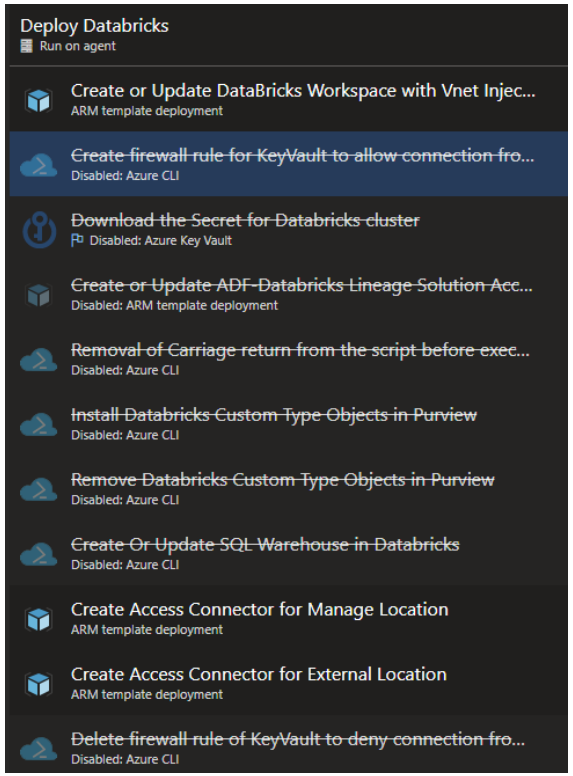


Example of Clinical Model Asset (SDTM Model)

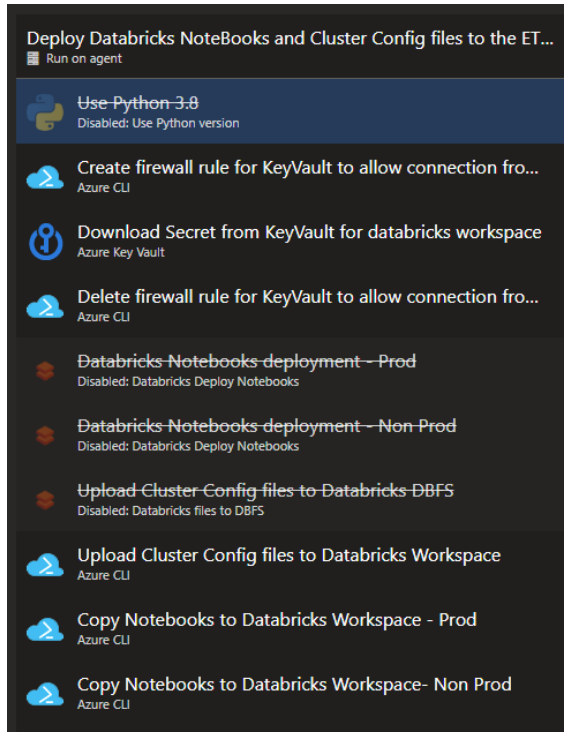


ALYCE Platform Automation Blueprint

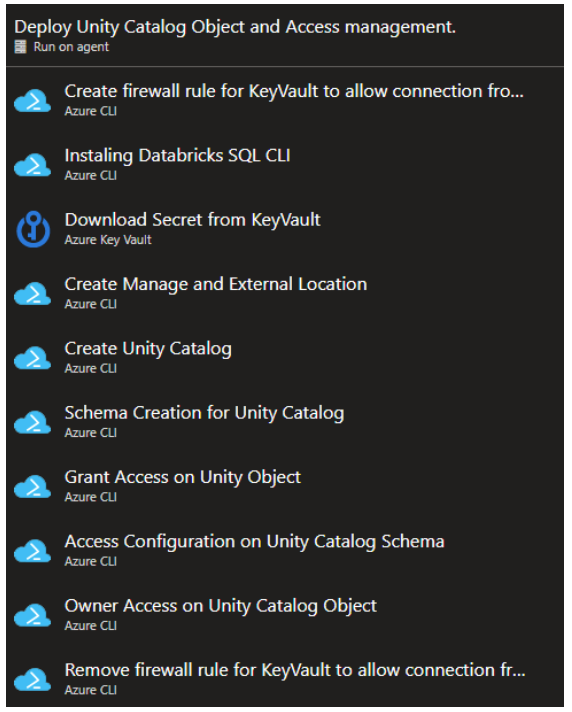
Infrastructure Pipeline



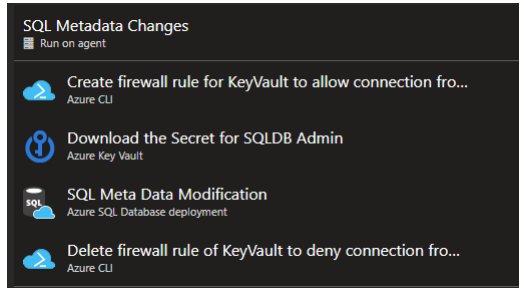
Data Pipeline

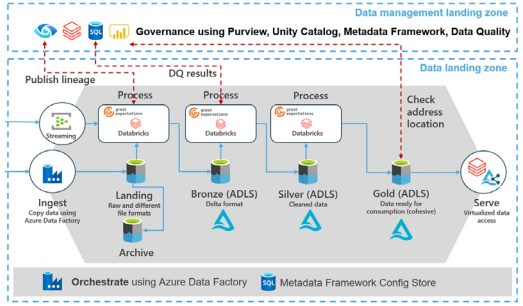


Access Pipeline



Operational Change Pipeline





ALYCE in a Nutshell

Management summary

Revolutionize (Clinical) Data Management: Shatter Data Silos and Boost Medical Review Speed by 10x and beyond.

Unleash the power of your R&D and organizational data by dismantling data silos and consolidating it using the Databricks Lakehouse platform. Witness a remarkable 10x acceleration in Medical Review efficiency, enabling faster, data-driven decisions.

Elevate your organization with the advanced Databricks Data Intelligence Platform, streamlining data analytics and decision-making processes.

Utilize the same data assets gathered through your comprehensive data pipelines for consistent, globally compliant medical reporting. Transform your organization's data capabilities and secure a competitive advantage in the rapidly evolving medical industry. Next stop will be Clinical Data Repository (CDR).

DATA+AI SUMMIT 2024

Thank you!