

RUN PYTHON SCRIPT IN THE CLOUD

Technical Approval Design – Phase 1 Deployment

Janua 2022



Purpose & Context

Initial pass and insight to the TDA

Purpose

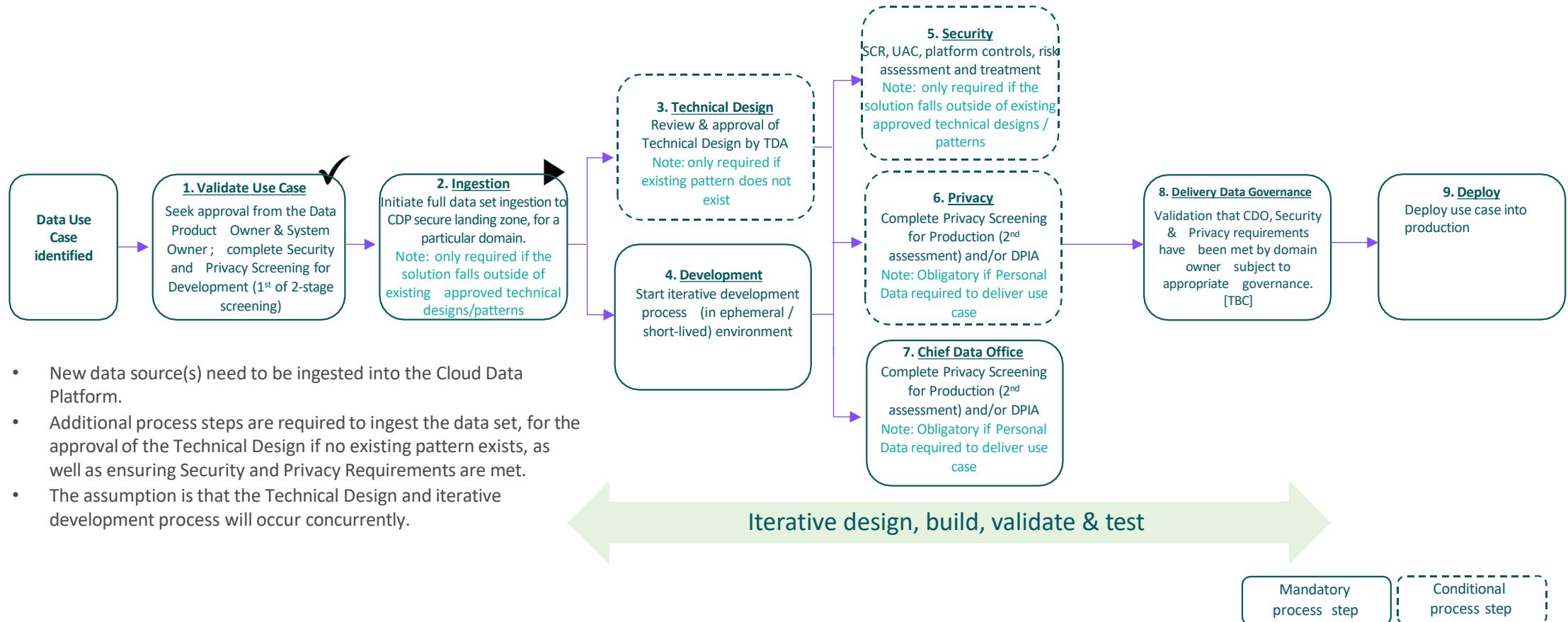
- To give initial pass and insight to the TDA before work starts. This not for final approval. Will come back to the TDA if changes and for final approval.
- The outlined principles have already been regularly discussed with Head of Architecture, Head of Data Science and Head of Platform.

Context

- Aligning the data on boarding with the new data journey process, which is working progress
- Phase 1 capabilities
- The phase 1 over laid the reference architecture and is only a sub part of the MVP this time

Onboarding Data to the Cloud

Data use journey process (Working Progress)



* Please note data usage refers to data required to satisfy a specific business use case; and is different to data access. A user might already have access to data, but not necessarily the approval to use the data for a new or different purpose.

Phase 1 – Run a python script in containers

The practical case will be for pre-processing journey data

Use Case Summary

The aim is not to bring in data at this phase, but to use the compiled text file that currently stored or even just use some mock data at this stage.

As the first phase is to have a pattern and process to run python scripts in containers that can scale in any cloud provider. Though this pattern won't be restricted to python scripts but anything that runs in a docker container (Python, .Net Core , Java, Nodejs, etc)

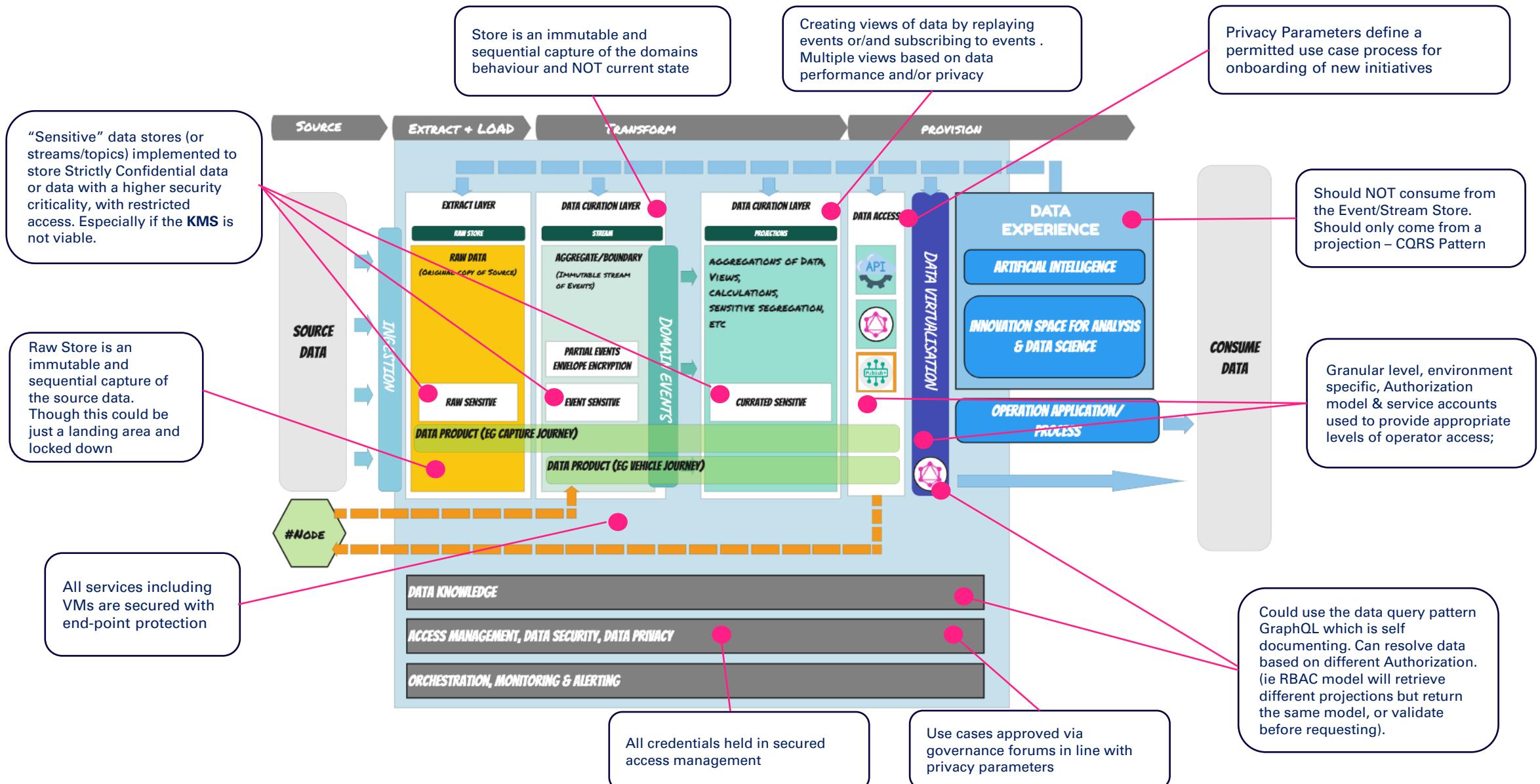
Phase Capabilities

- A pattern/process of deploying and running python scripts in containers
- It's not necessarily restricted to Python
- It's not necessarily restricted to AWS
- Branching Strategy, along with a Release Process
- Stretch: Python Script can publish to a Bus or pub/sub.

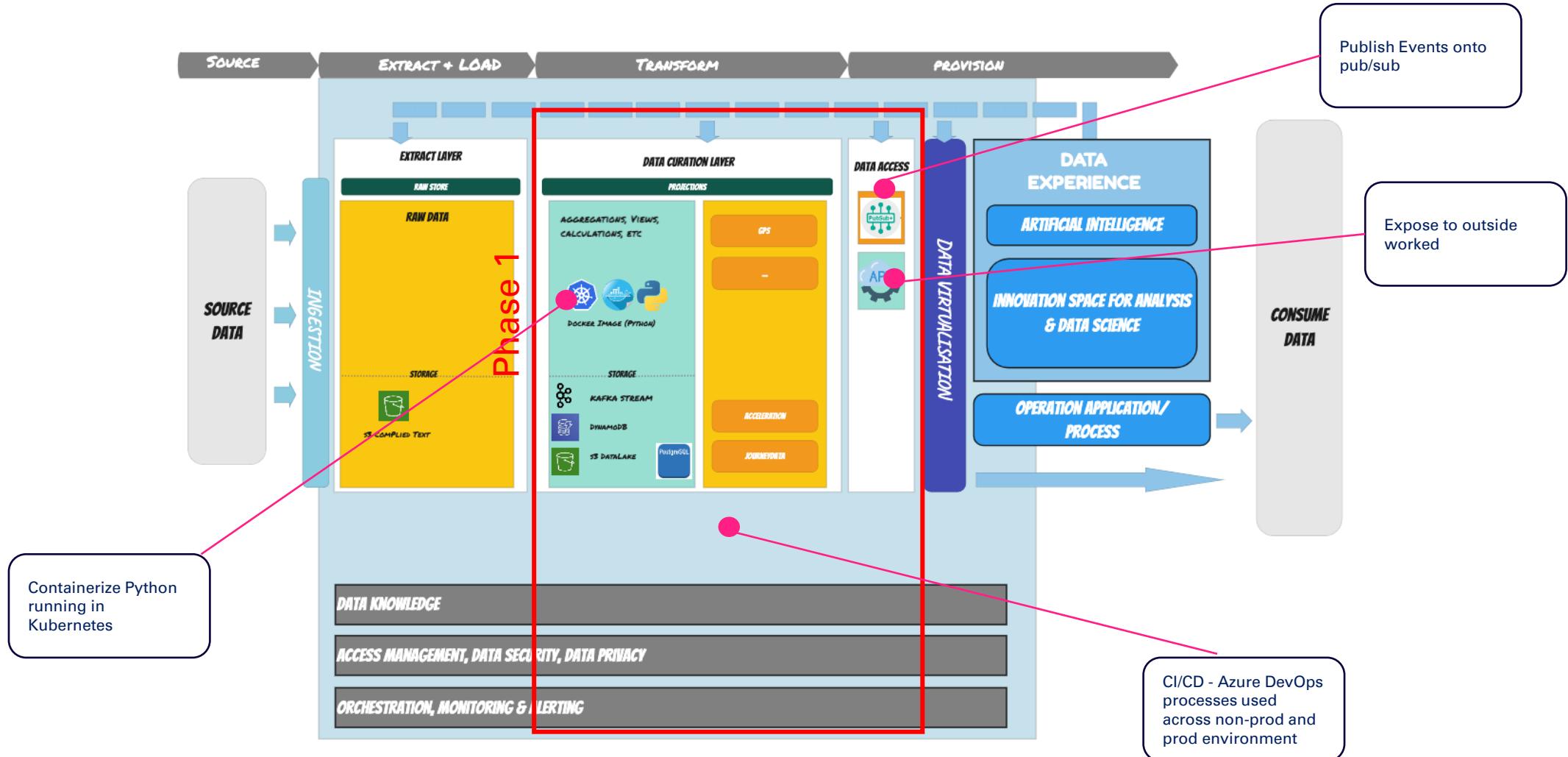
Data Privacy and Security Considerations

- At this stage NO one will have access to the data at this stage.
- Will work closely with security and privacy teams to define an appropriate treatment method to allow for the safe storage and use of former customer data (e.g. in a depersonalised way) if there is any. Though for this phase will try to avoid any sensitive or GDPR data.

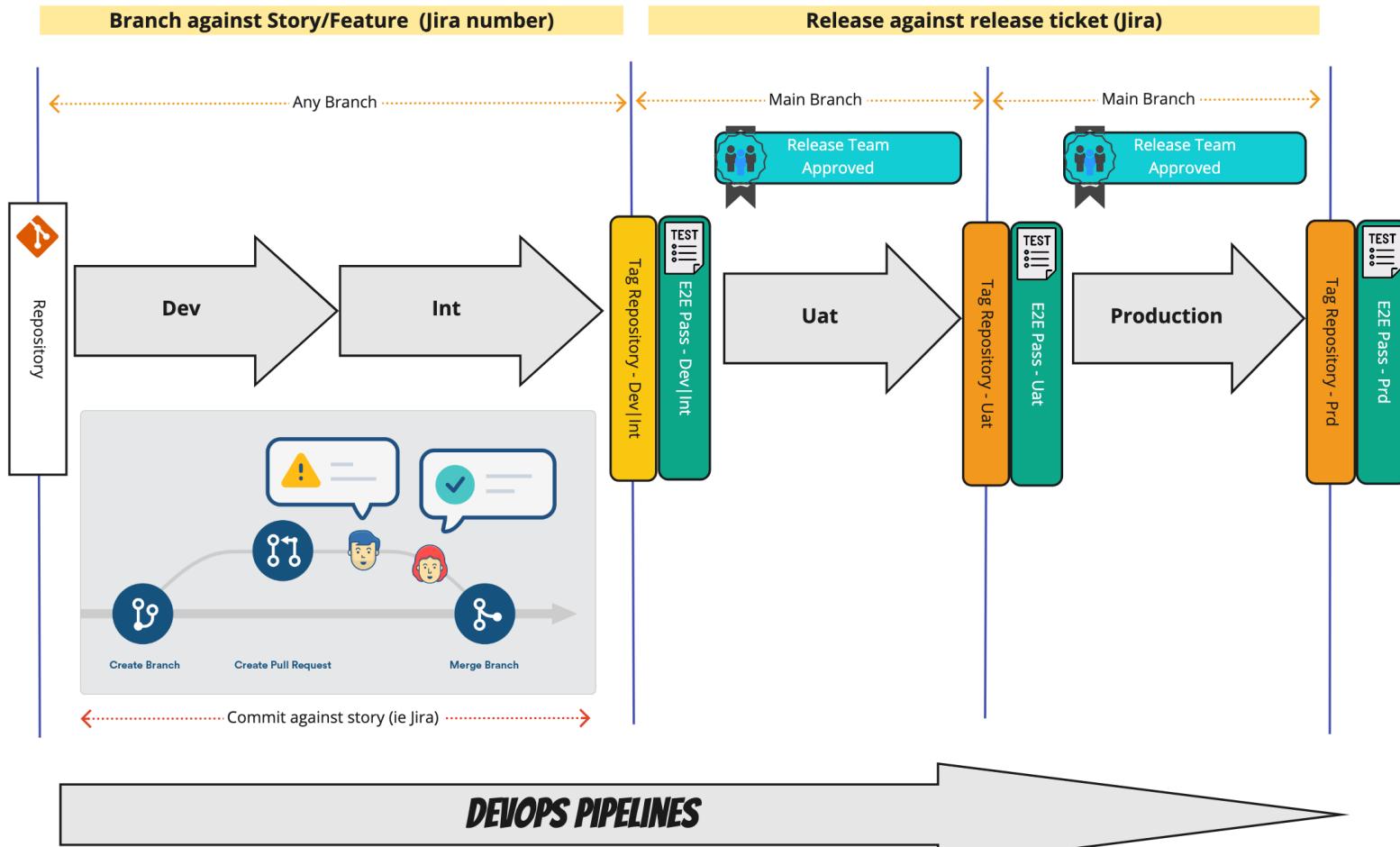
Reference Architecture – Data Product Controls



Reference Architecture – with Technology over laid



Azure DevOps Pipeline



Appendix

Reference Architecture – With KMS

