

**OLF** AI & DATA



The  
**EGERIA**  
Metadata Show

# THE VALUE OF EGERIA AND OPEN METADATA

Mandy Chessell CBE FREng  
Egeria Open Source Project Lead

# Webinar Schedule 2021

Date	time	Title	Description	Presenter
<b>Date change</b> <b>6th 13th</b> <b>September</b> <b>2021</b>	15:00 UTC	<b>Visualising a metadata ecosystem</b>	The session is for people looking to understand the metadata across their ecosystem in terms of the Egeria open types and instances using visualisations in the Egeria React User Interface. Understanding the types is important knowledge when developing connectors and new APIs like OMAS's. This session will also show how metadata instances can be explored at a low level. This will be contrasted with an exploration of semantic data that is based on the Subject Area Open Metadata Access Service (OMAS).  Zoom Conference <a href="https://zoom.us/j/523629111">https://zoom.us/j/523629111</a>	<b>David Radley</b>
<b>4th</b> <b>October</b> <b>2021</b>	15:00 UTC	<b>The Value Egeria brings to an organisation.</b>	This session is for people wanting to understand the value of Egeria in enabling data centric, metadata driven integration. The session will start with the core Egeria constructs, including entities, explaining the principles behind why they are as they are. The session will go through the layers and aspects of the Egeria architecture, at each stage talking about the applicability to solving real world problems. By the end of the session you should have awareness of the parts of Egeria at a high level, why they have been implemented as they are and the value that each of the pieces bring.  Zoom Conference <a href="https://zoom.us/j/523629111">https://zoom.us/j/523629111</a>	<b>Mandy Chessell</b>
<b>8th</b> <b>November</b> <b>2021</b>	15:00 UTC	<b>Open lineage</b>	This session will describe how to set up an Egeria production system where lineage can be visualised based on metadata from multiple sources. It will explain the Open Lineage standard and the Open Lineage services - two different things using the words "Open Lineage".  This will include scenarios: <ul style="list-style-type: none"> <li>• metadata accessed using the OMRS repository connector proxy to a proprietary metadata repository</li> <li>• A scenario with all open source content and no proprietary code.</li> <li>• Lineage capture in various use cases</li> <li>• Consolidated view</li> </ul>	<b>ING Bank (TBA) and Mandy Chessell</b>
<b>6th</b> <b>December</b> <b>2021</b>	15:00 UTC	<b>Kubernetes operators and Egeria</b>	This session will cover how easy it is to run Egeria in Kubernetes and how the Egeria Kubernetes operator can be used to manage Egeria in a Kubernetes environment.	<b>Nigel Jones</b>

# Webinar Schedule 2022

<b>10th January 2022</b>	15:00 UTC	<b>What next after you have built a catalog. Part 1: the journey</b>	Journey from manual cataloging, to using automated integration and templating. Metadata discovery and stewardship will be covered here also as well as metadata deduplication.	<b>Mandy Chessell</b>
<b>7th February 2022</b>	15:00 UTC	<b>Time Travelling with Egeria</b>	Every wanted to know what the state of your metadata was at some specific time in the past? This session will introduce the Crux open metadata repository that supports these historical metadata queries.	<b>Chris Grote</b>
<b>7th March 2022</b>	15:00 UTC	<b>How to build a repository connector</b>	Every wanted to build an OMRS repository connector? This session will take you though what the considerations are and you need to do. It will show how to create the simplest "Hello World" connector.	<b>Chris Grote</b>
<b>4th April 2022</b>	15:00 UTC	<b>How to build an integration connector</b>	This session covers how to extend Egeria's automated cataloguing to include metadata from a new technology. It describes how automated cataloguing works and the role of the integration connector. It covers the design of the integration connector using examples to illustrate the different approaches and their benefits and and challenges. It shows how to set up a project for a new connector, how to build and package it and finally it shows the new connector running in Egeria.	<b>Mandy Chessell</b>
<b>9th May 2022</b>	15:00 UTC	<b>What next after you have built a catalog. Part 2: making it real</b>	Making it real in your organisation. What are the infrastructure and culture changes that drive successful management of data? How can you use Egeria to design your governance organization and drive the desired cultural changes and behaviour.	<b>Mandy Chessell</b>
<b>6th June 2022</b>	15:00 UTC	<b>What next after you have built a catalog. Part 3: strategic view</b>	Becoming more mature with data governance, eventually enabling data citizenship. Data citizenship within an organization supports people who need access to data to be effective in their role. This data access needs to be embedded in the tools that they use and flexible enough to meet the changing challenges they face day to day. Since the data landscape is also changing, the tools that these people use need to be integrated with the data catalog.	<b>Mandy Chessell</b>

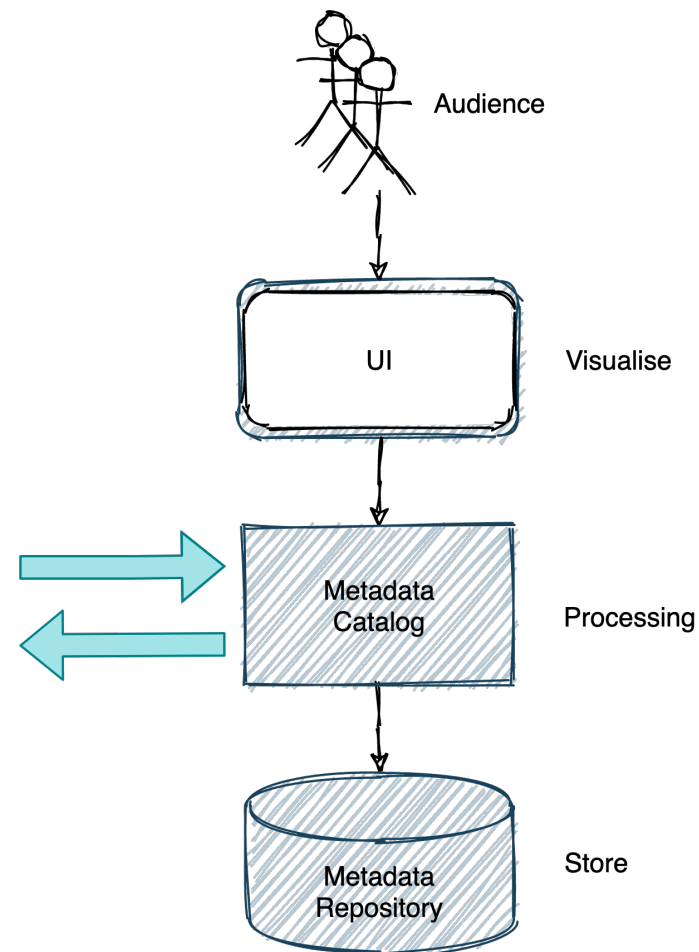
# Today's Agenda

- Supporting a search across multiple catalogs
  - Mechanisms
  - Metadata types
- Supporting the request for more data
  - Additional metadata types
- Getting started



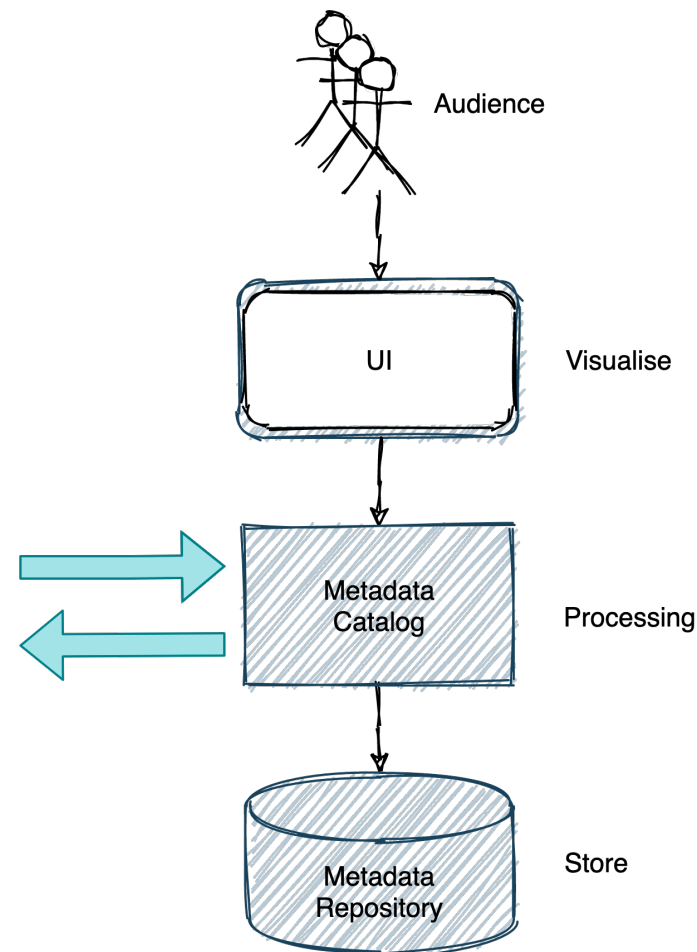
# Supporting metadata queries

- Typically metadata catalogs can only display and maintain metadata that is stored in its repository



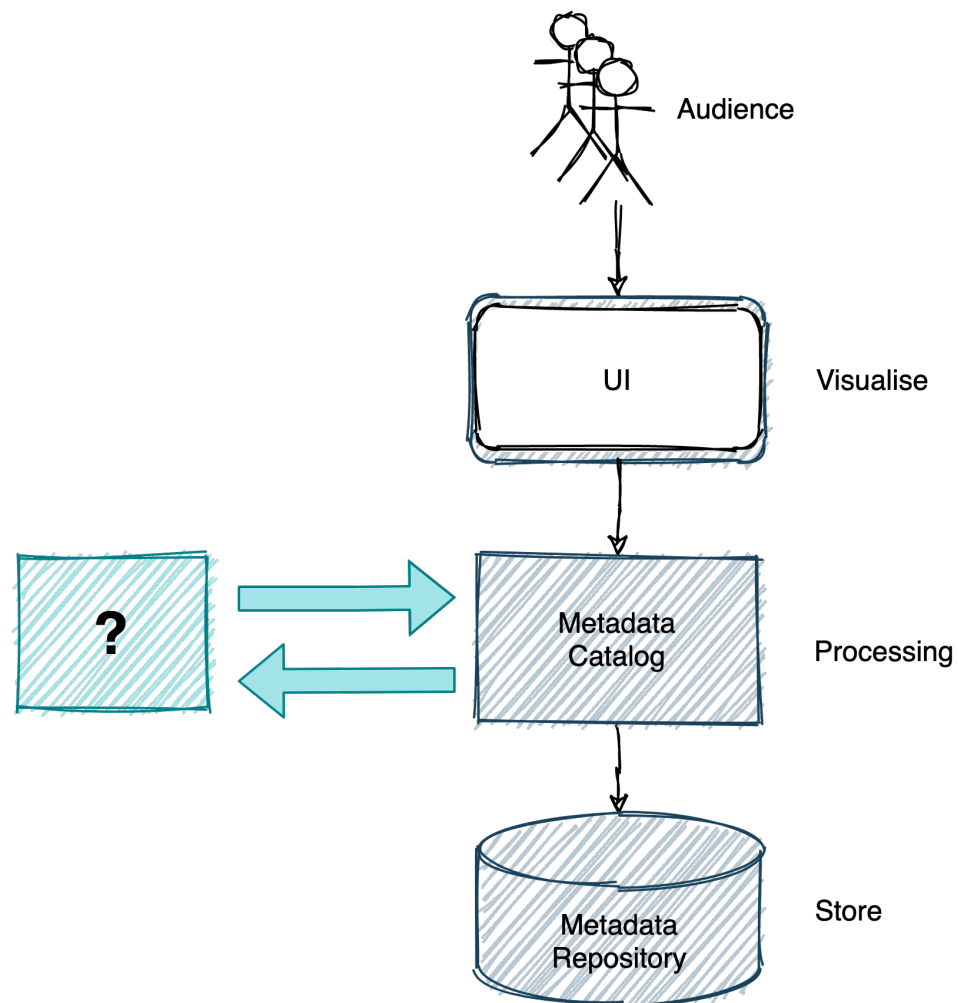
## Evaluating the roles of a metadata catalog in the open metadata ecosystem

- What metadata is needed by the audience?
- What types of metadata can be stored?
- What value-add services does the repository provide?
  - onboarding and maintenance of metadata
- How sophisticated is the catalog API and events

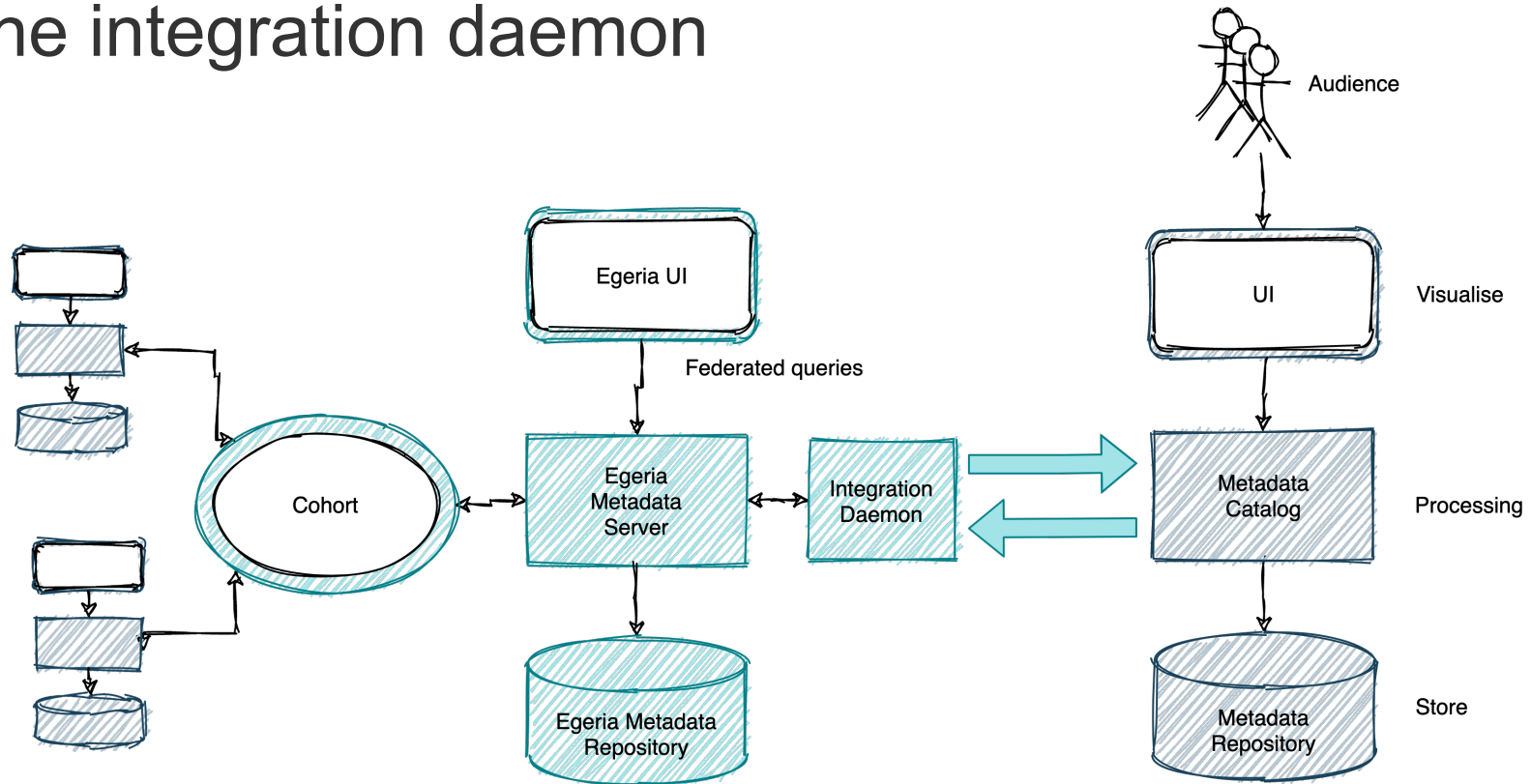


# Choosing the integration style

- How sophisticated is the catalog API and events?
- How available is the metadata catalog?
- Does the metadata catalog have the capacity to support queries?



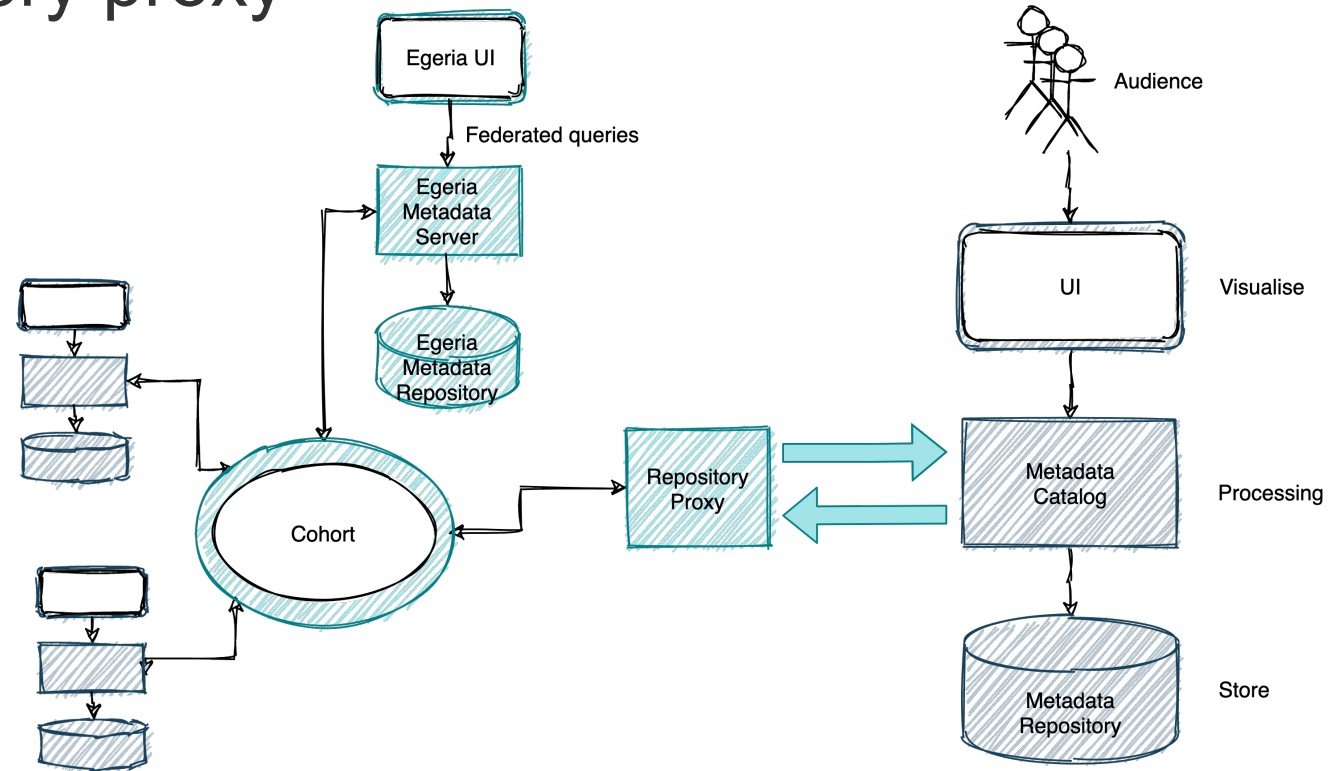
# Using the integration daemon



Egeria manages the home copy of the metadata.

- Available even if metadata catalog is not
- Minimal additional load on the metadata catalog
- Metadata passed into catalog has benefit of Egeria's metadata governance (eg deduplication)

# Using the repository proxy

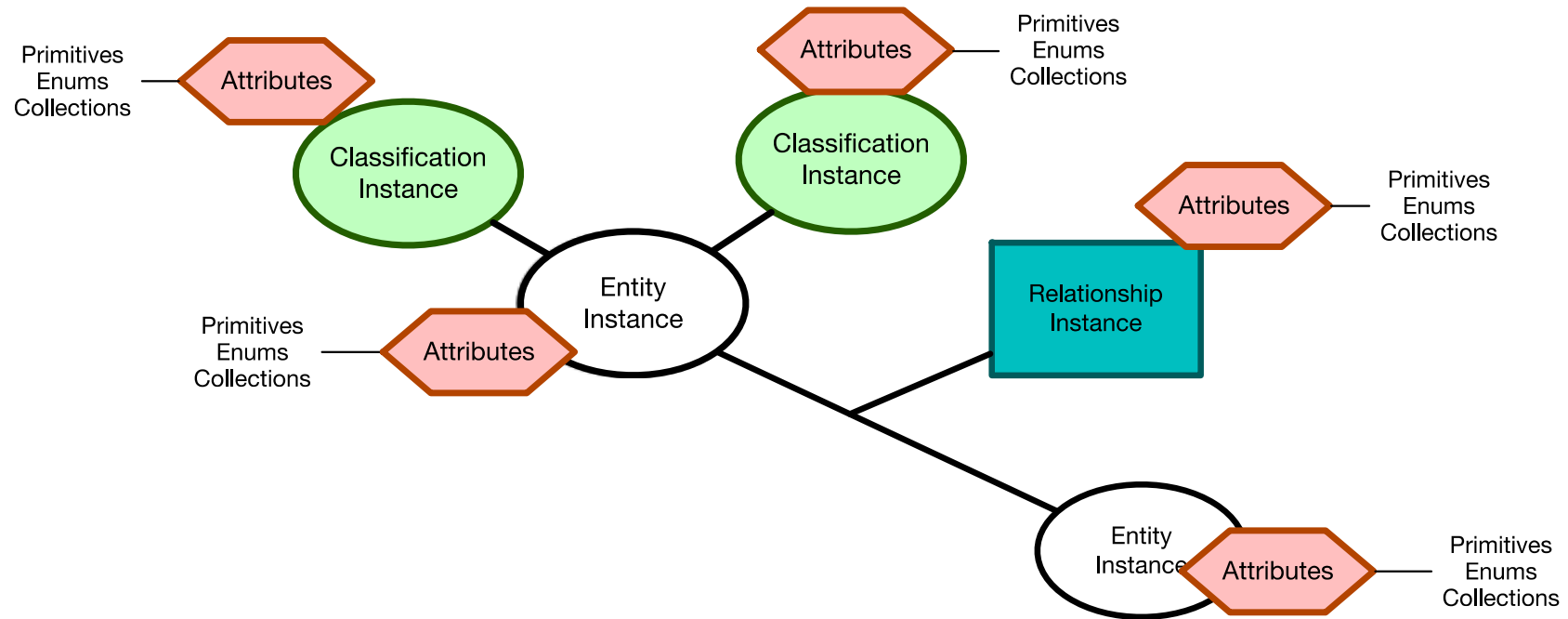


Metadata catalog connects directly to cohort via a repository proxy

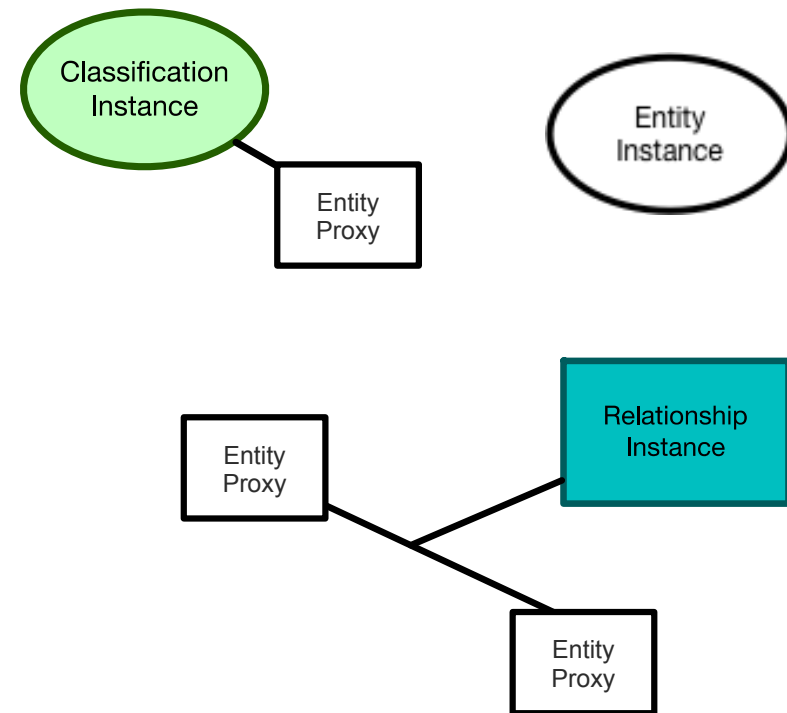
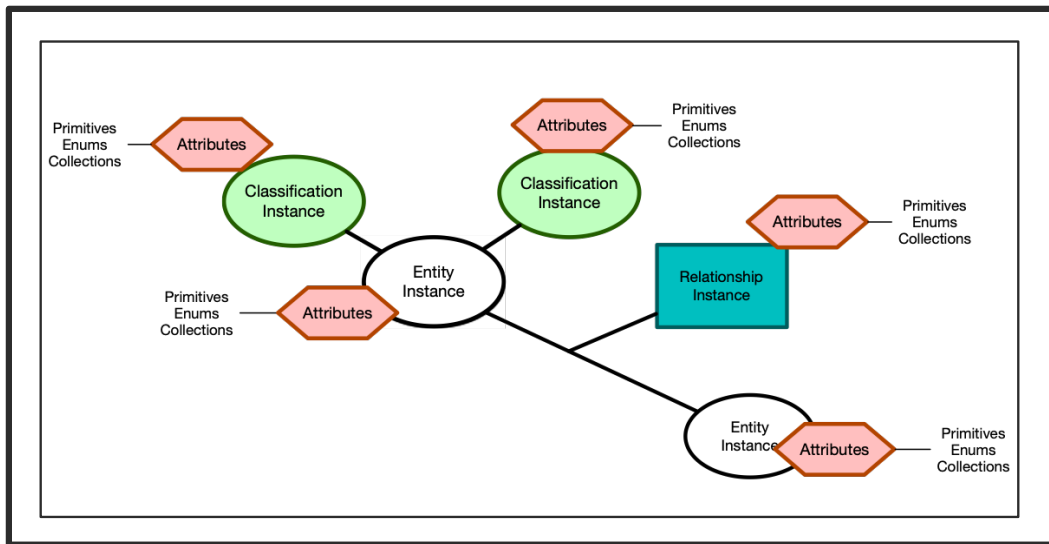
- Receives all metadata elements
- Must support query interface for its metadata to be visible to federated queries
- (If no query interface then need to ensure all of its metadata is copied into an Egeria metadata server)



# Instance representations in the metadata in the cohort

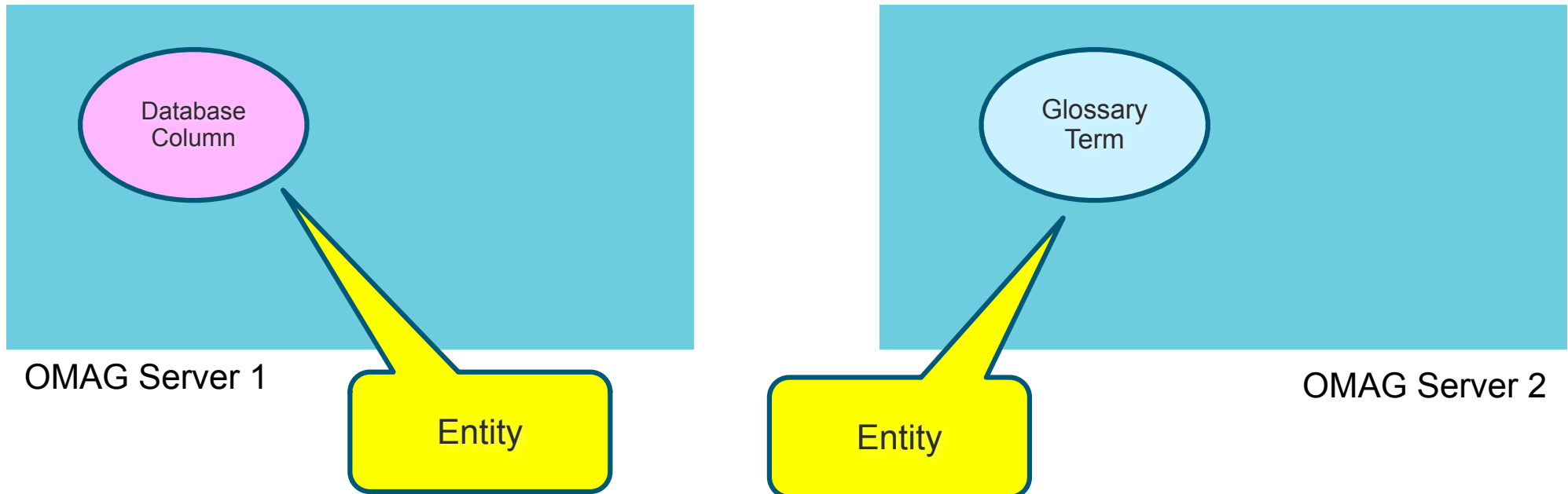


# Instance representations in the metadata in the cohort

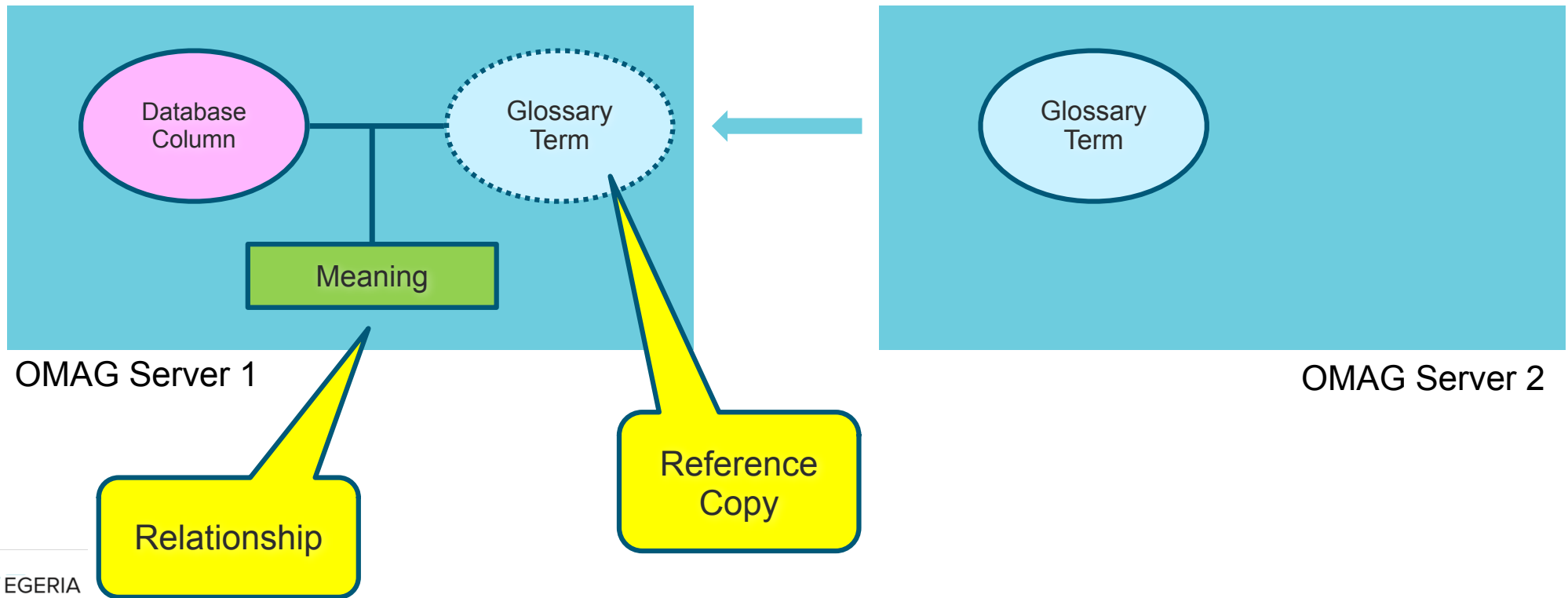


Graph fragments for APIs, events, archives and repositories

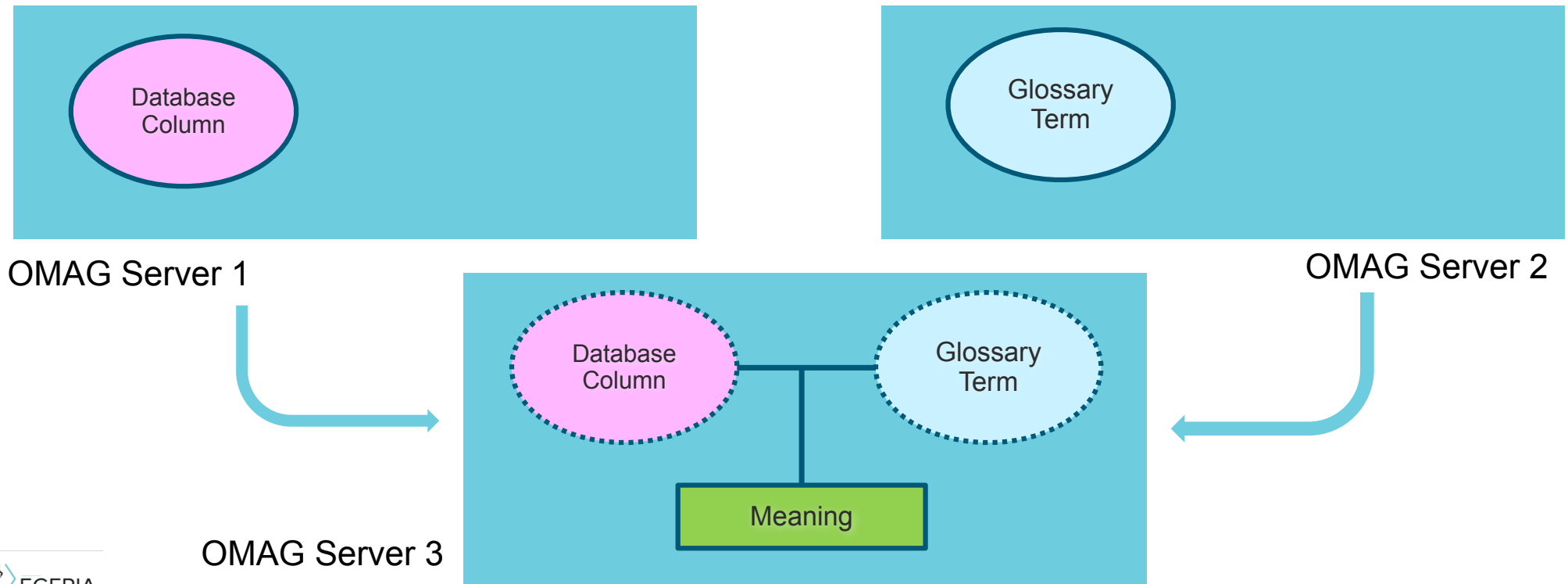
# Linking metadata across the cohort



# Linking metadata across the cohort

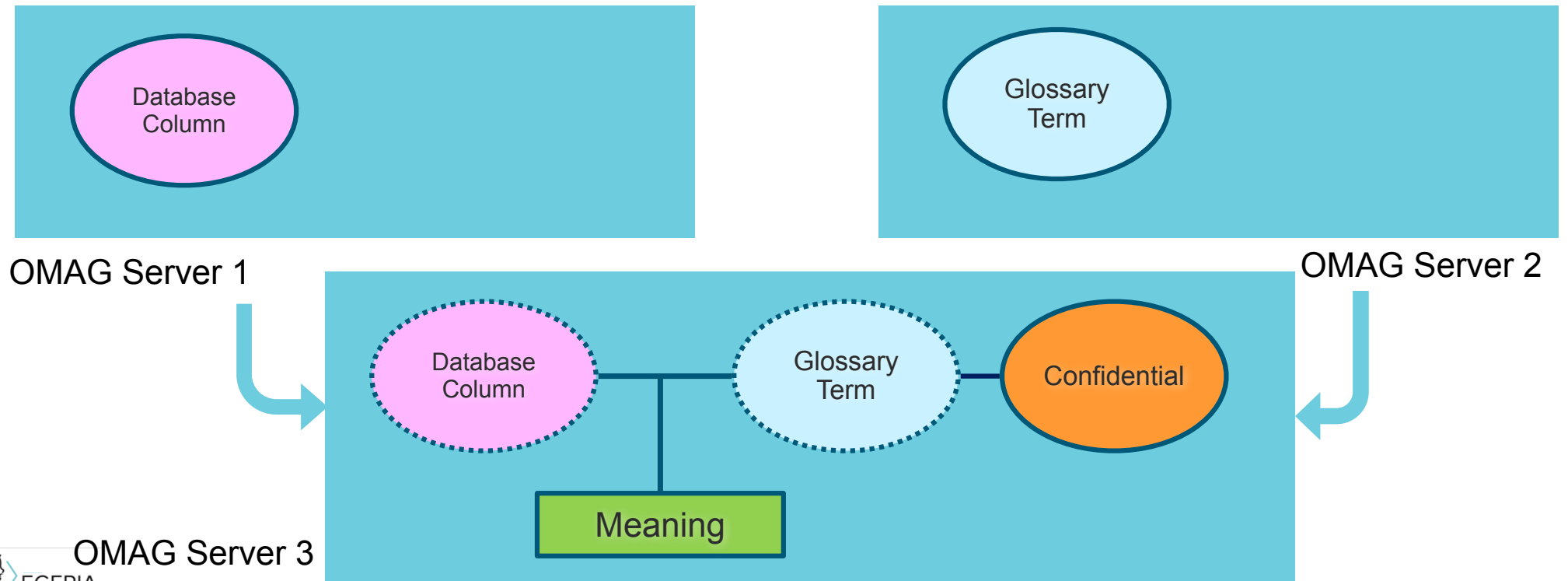


# Linking metadata across the cohort

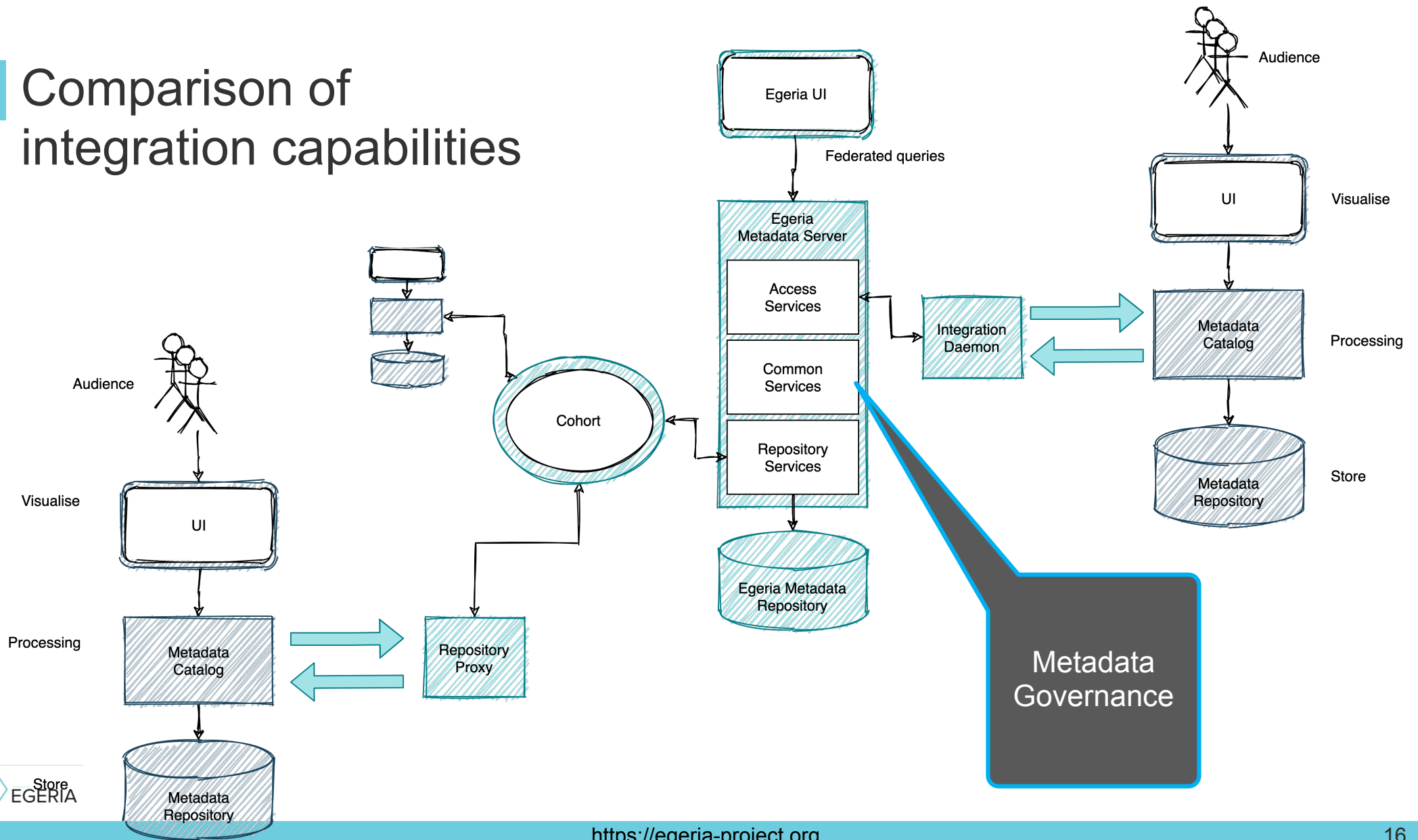




# Linking metadata across the cohort

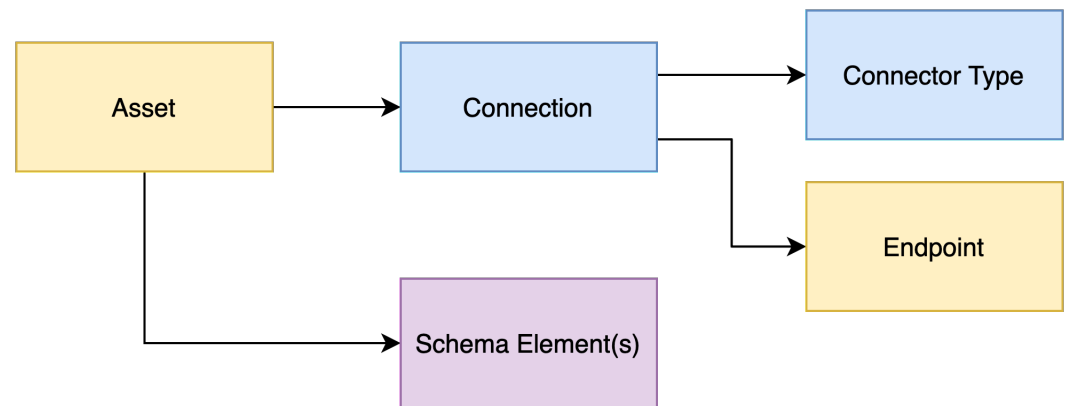


# Comparison of integration capabilities

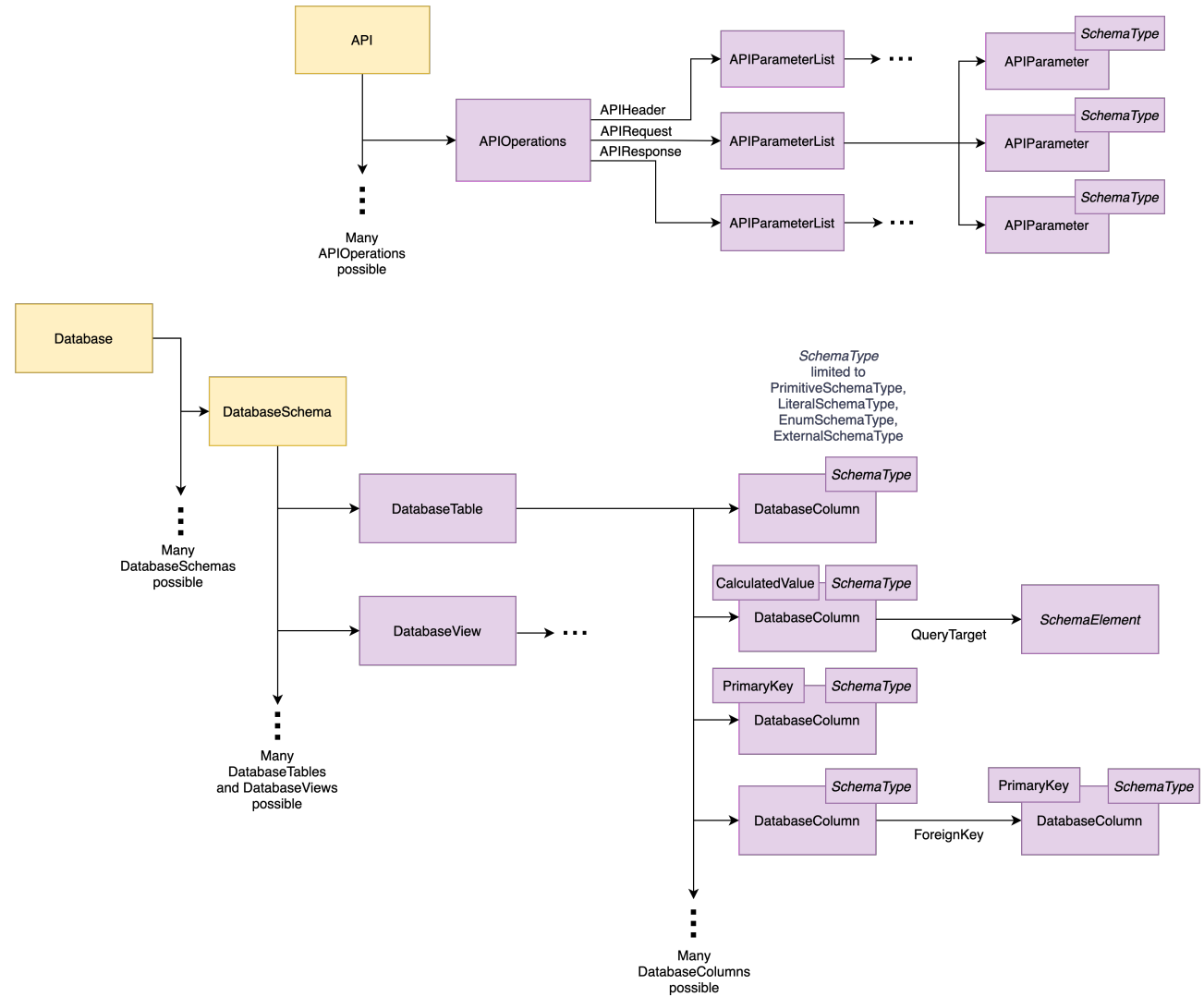
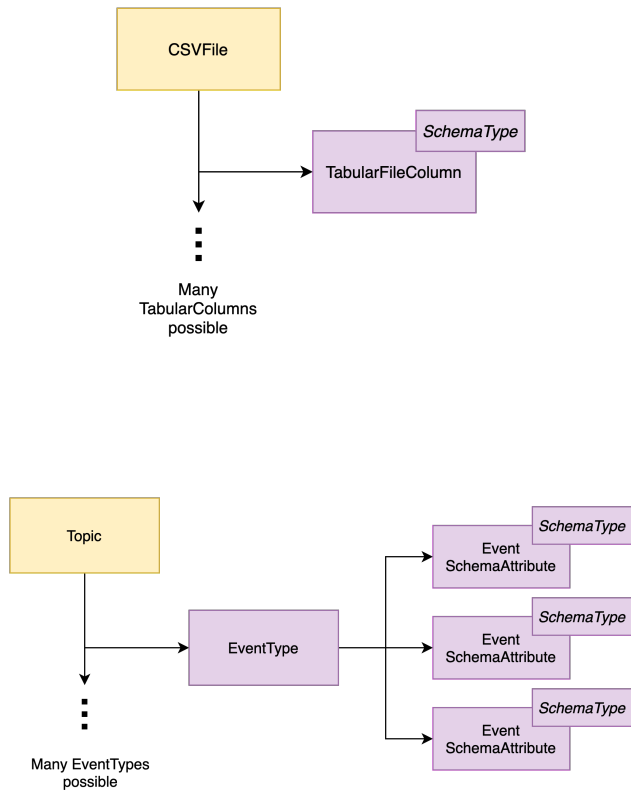


# Basic Model - Technical Metadata

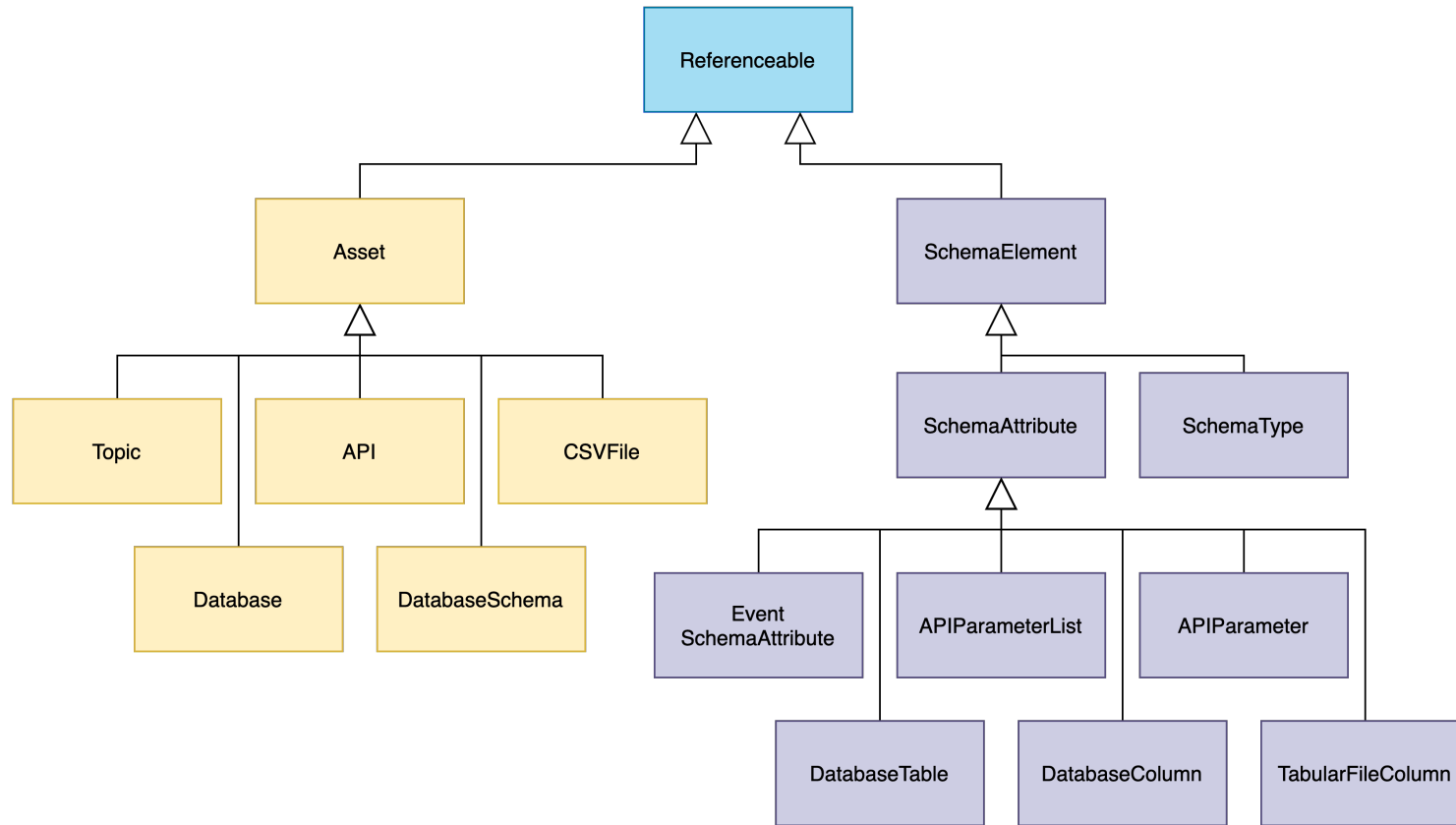
- Asset - resource to govern
- Connection - access information to real resource
  - Connector type - connector implementation to use
  - Endpoint - network location
- Schema elements - structure of the data



# Schema Structures



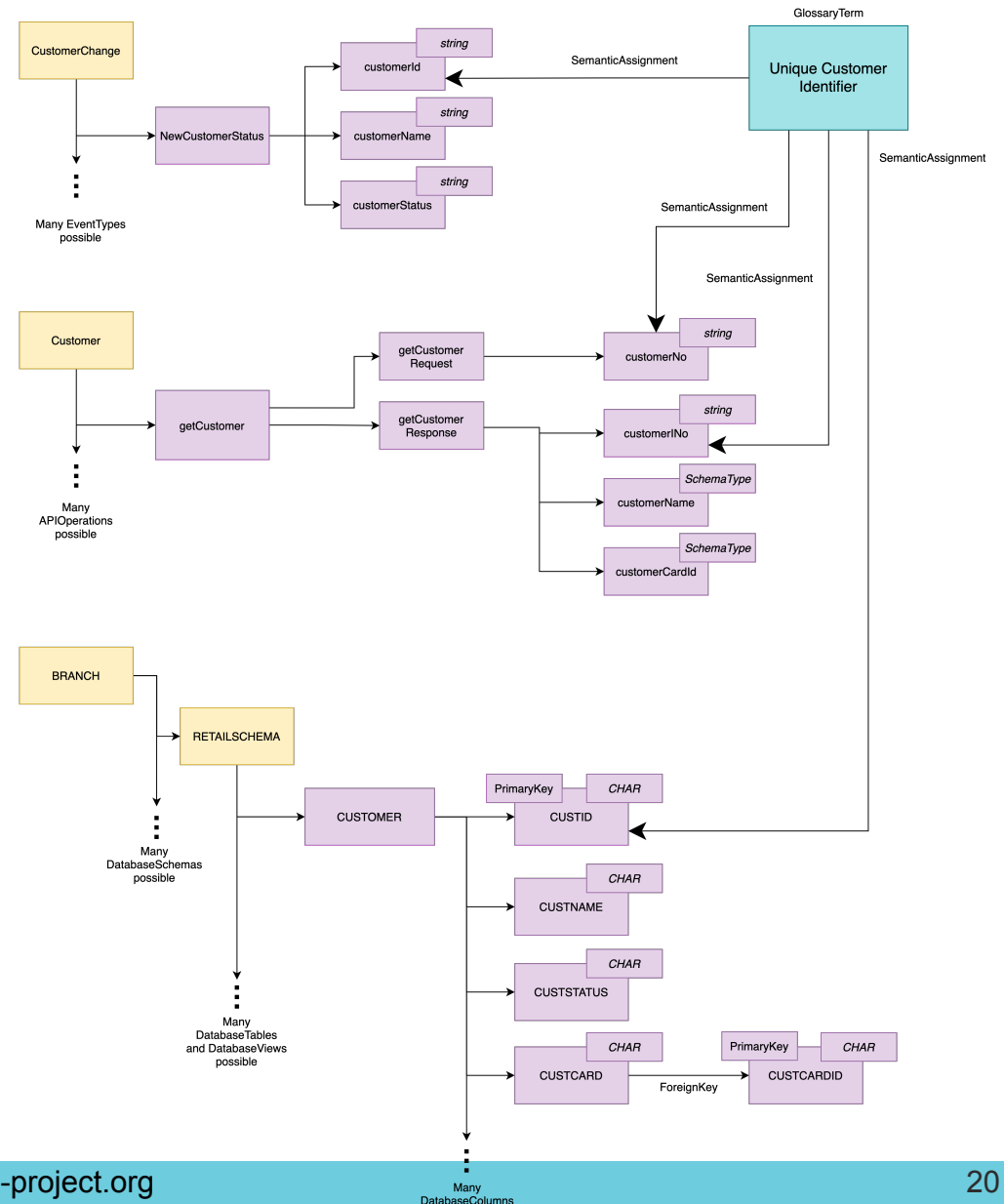
# Type Inheritance





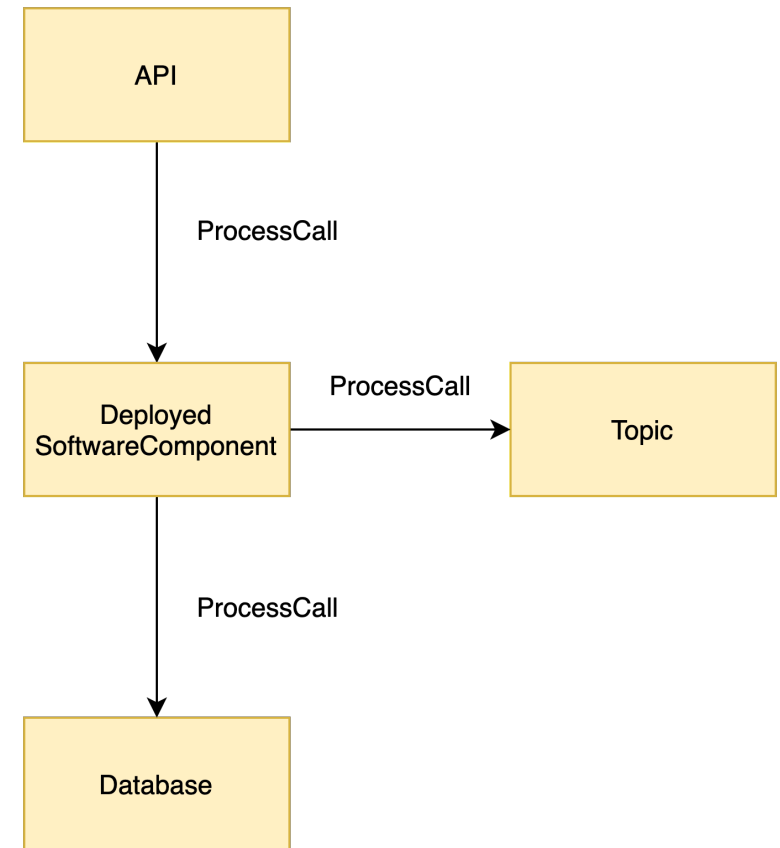
# Building a data map

- Using semantic assignment
- Additional options
  - Search keywords
  - Informal Tags
  - Reference Data Value
  - Governance Classifications
  - Location
  - Lineage
  - *more ...*

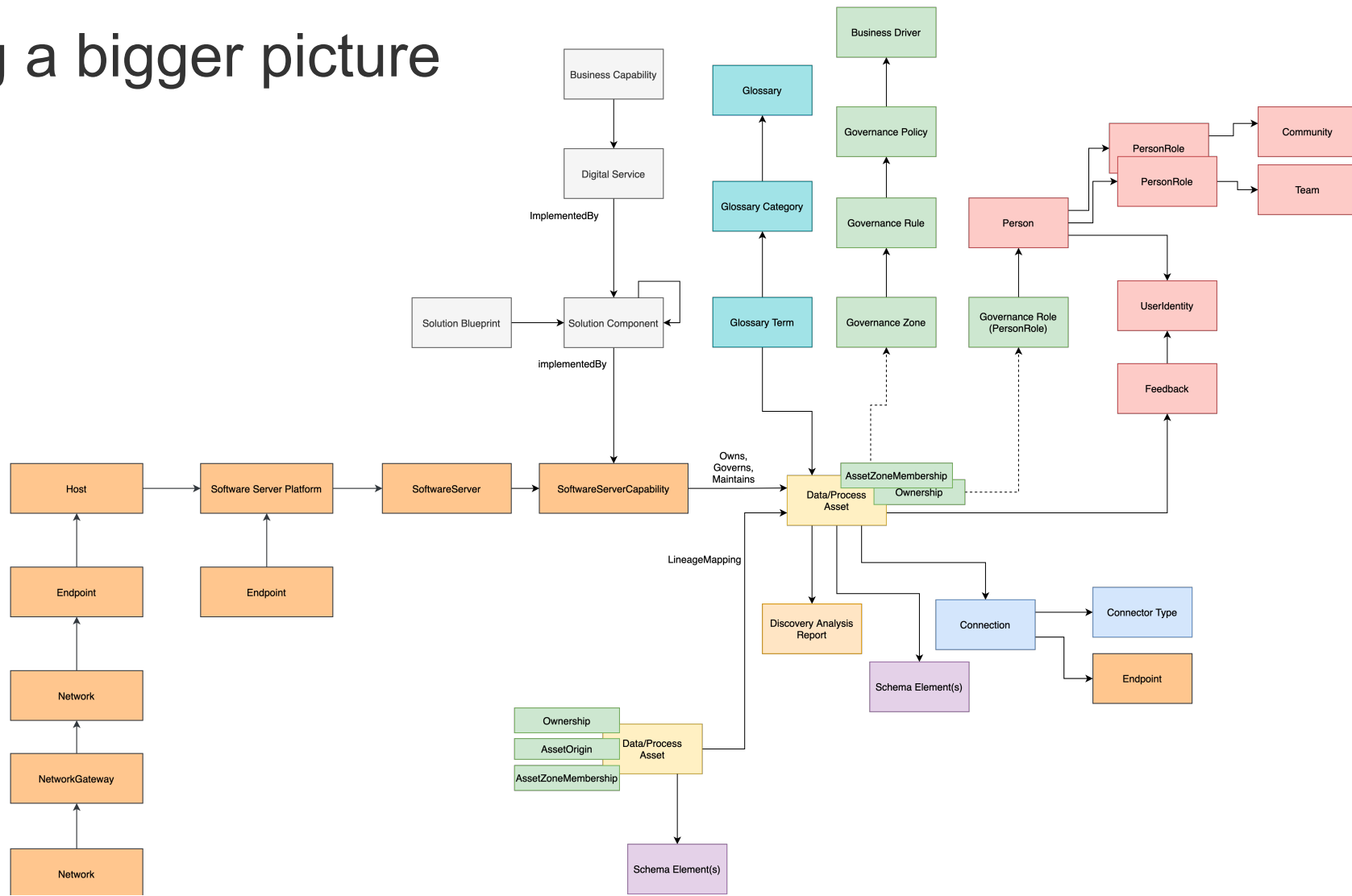


# Stitching

- There are a variety of options for linking APIs, events and data together.
  - Process calls
  - Data flows
  - Control flows
  - Lineage mappings

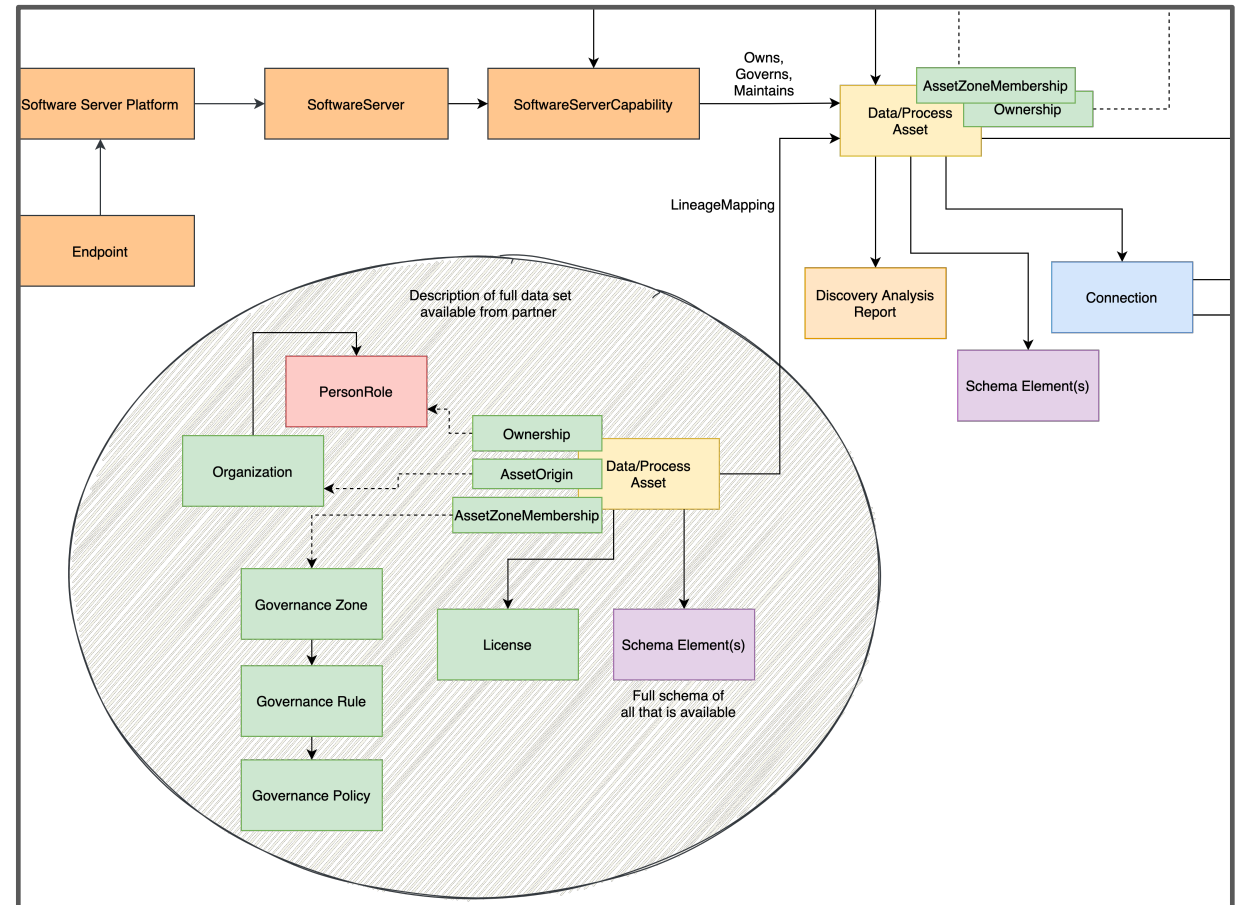


# Building a bigger picture



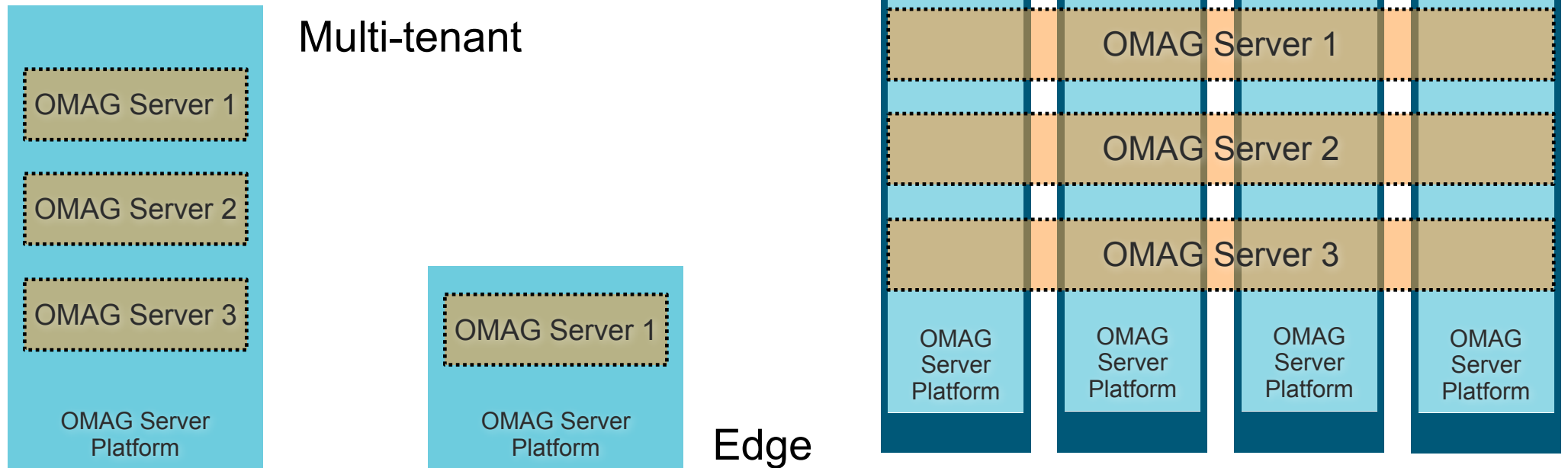
# Adding more information

- Lineage shows the source
- Attachments to the source define how to get more data



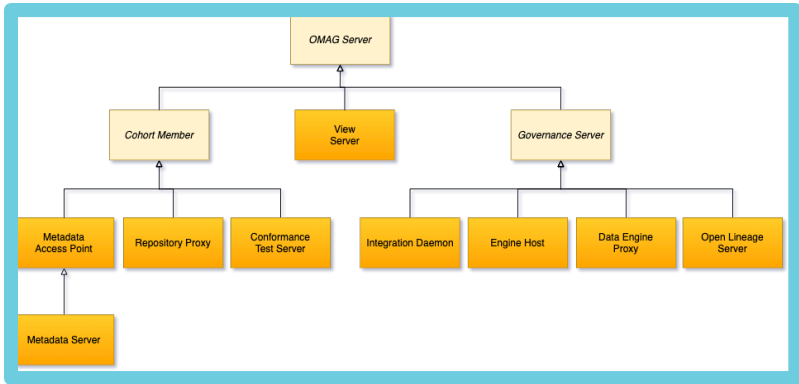
# Egeria's OMAG Server Platform

Open Metadata and Governance

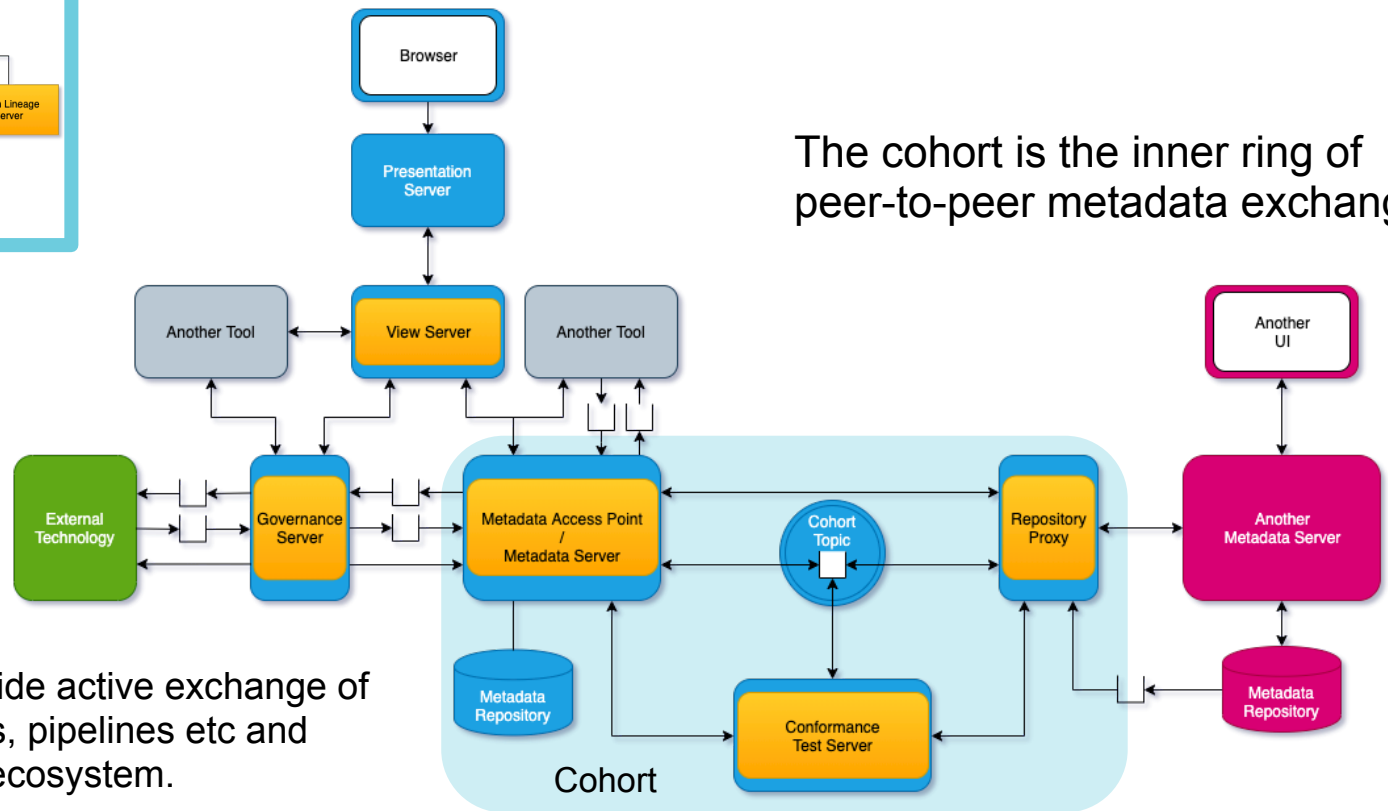




# Egeria's Open Metadata and Governance (OMAG) Servers



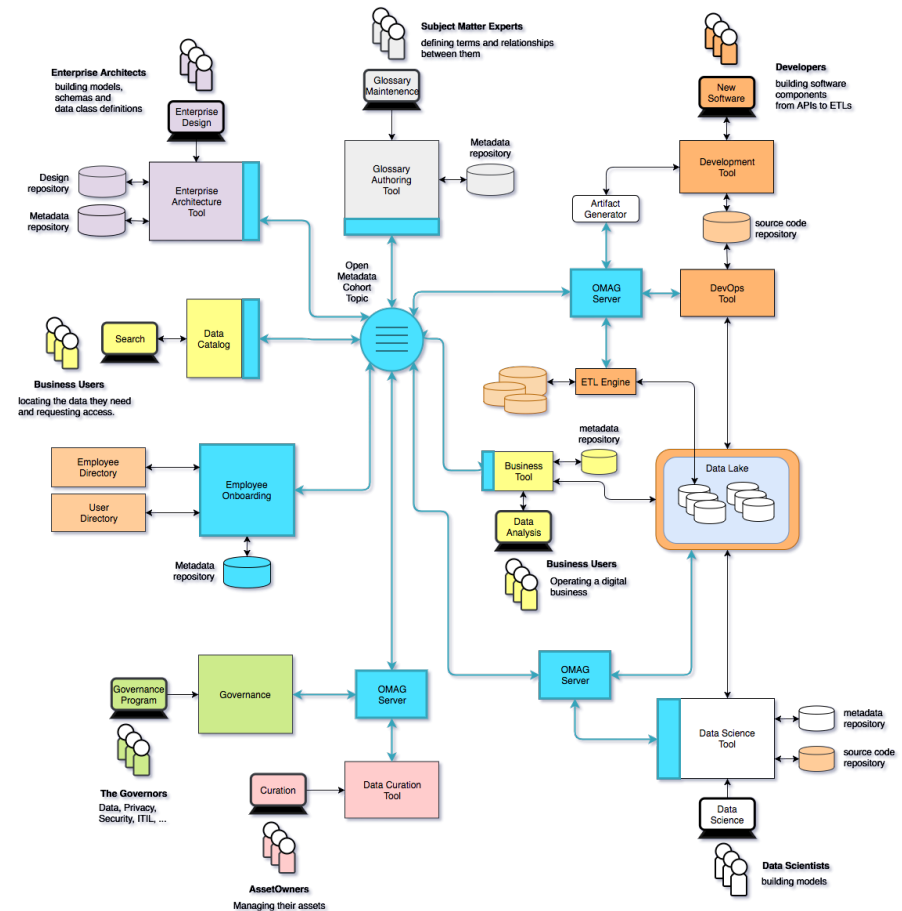
The cohort is the inner ring of peer-to-peer metadata exchange.



The governance servers provide active exchange of metadata with tools, platforms, pipelines etc and governance of the metadata ecosystem.

# Using Egeria ...

- Eases the cost of metadata integration through
  - Comprehensive standards and libraries.
  - Active vendor recruitment program.
- Provides direct support to many governance roles, filling the gaps between function offered through commercial tools.
- Provides best practices and content packs to accelerate an organization's journey to becoming data driven.



# Open forum



# Webinar Schedule 2021

Date	time	Title	Description	Presenter
<b>Date change</b> <b>6th 13th</b> <b>September</b> <b>2021</b>	15:00 UTC	<b>Visualising a metadata ecosystem</b>	The session is for people looking to understand the metadata across their ecosystem in terms of the Egeria open types and instances using visualisations in the Egeria React User Interface. Understanding the types is important knowledge when developing connectors and new APIs like OMAS's. This session will also show how metadata instances can be explored at a low level. This will be contrasted with an exploration of semantic data that is based on the Subject Area Open Metadata Access Service (OMAS).  Zoom Conference <a href="https://zoom.us/j/523629111">https://zoom.us/j/523629111</a>	<b>David Radley</b>
<b>4th</b> <b>October</b> <b>2021</b>	15:00 UTC	<b>The Value Egeria brings to an organisation.</b>	This session is for people wanting to understand the value of Egeria in enabling data centric, metadata driven integration. The session will start with the core Egeria constructs, including entities, explaining the principles behind why they are as they are. The session will go through the layers and aspects of the Egeria architecture, at each stage talking about the applicability to solving real world problems. By the end of the session you should have awareness of the parts of Egeria at a high level, why they have been implemented as they are and the value that each of the pieces bring.  Zoom Conference <a href="https://zoom.us/j/523629111">https://zoom.us/j/523629111</a>	<b>Mandy Chessell</b>
<b>8th</b> <b>November</b> <b>2021</b>	15:00 UTC	<b>Open lineage</b>	This session will describe how to set up an Egeria production system where lineage can be visualised based on metadata from multiple sources. It will explain the Open Lineage standard and the Open Lineage services - two different things using the words "Open Lineage".  This will include scenarios: <ul style="list-style-type: none"> <li>• metadata accessed using the OMRS repository connector proxy to a proprietary metadata repository</li> <li>• A scenario with all open source content and no proprietary code.</li> <li>• Lineage capture in various use cases</li> <li>• Consolidated view</li> </ul>	<b>ING Bank (TBA) and Mandy Chessell</b>
<b>6th</b> <b>December</b> <b>2021</b>	15:00 UTC	<b>Kubernetes operators and Egeria</b>	This session will cover how easy it is to run Egeria in Kubernetes and how the Egeria Kubernetes operator can be used to manage Egeria in a Kubernetes environment.	<b>Nigel Jones</b>

# THANK YOU!

**LinkedIn:** <http://www.linkedin.com/pub/mandy-chessell/22/897/a49>

**Slack:** <https://lfaifoundation.slack.com/archives/C01F40J2XA8>

**Email:** [egeria-technical-discuss@lists.lfaidata.foundation](mailto:egeria-technical-discuss@lists.lfaidata.foundation)

