

Egeria Webinar

USING INTEGRATION CONNECTORS

Mandy Chessell CBE FREng
Egeria Open Source Project Lead

Date	time	Title	Description	Presenter
January 2022	n/a	n/a	There is no webinar this month. The team is working on creating a new set of monthly Dojo sessions . The first of these sessions will be on the 17th of January 2022 10AM - 17:00PM (UTC) and will be on How to setup and run Egeria in different environments . This will include the Kubernetes environment. Zoom Conference https://zoom.us/j/523629111	
7th February 2022	15:00 UTC	Using an integration connector	This session covers how to use Integration connectors to connect technologies into Egeria..	Mandy Chessell
7th March 2022	15:00 UTC	How to build an integration connector	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 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.	Mandy Chessell
4th April 2022	15:00 UTC	Using a repository connector	This session covers how to use Repository Connectors to connect technologies into Egeria; focussing on XTDB (formerly known as crux). Every wanted to know what the state of your metadata was at some specific time in the past? This session will introduce the XTDB open metadata repository that supports these historical metadata queries.	Chris Grote

Coco Pharmaceuticals persona



Jules Keeper, CDO



Tessa Tube,
Chief Researcher



Faith Broker
Chief Privacy Officer



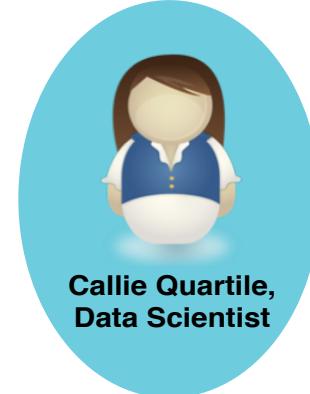
Nancy Noah
Cloud Specialist



Erin Overview,
Information Architect



Bob Nitter,
Integration Developer



Callie Quartile,
Data Scientist

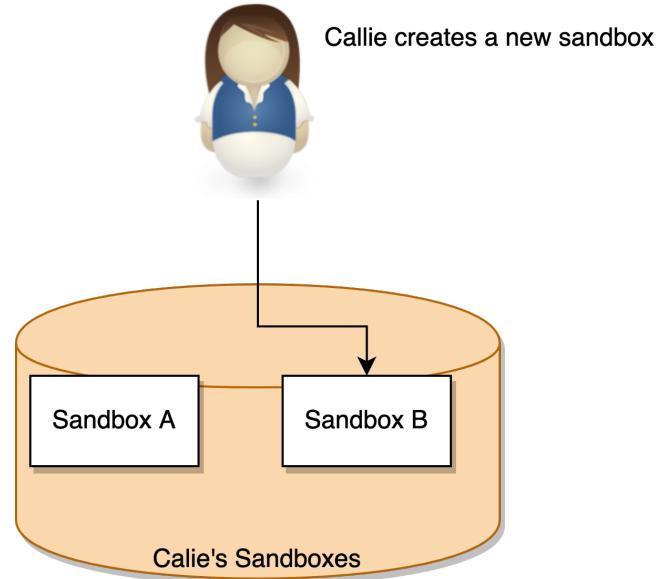


Gary Geeke
IT Infrastructure

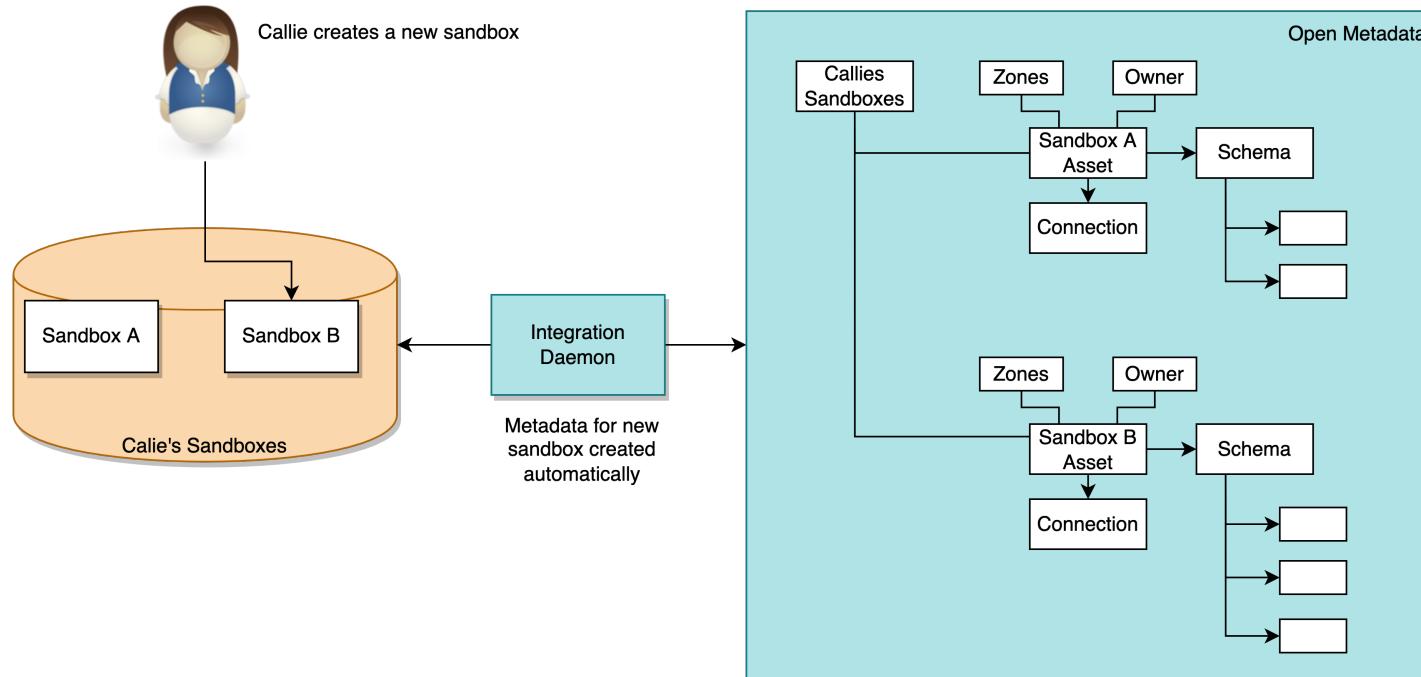
<https://odpi.github.io/data-governance/coco-pharmaceuticals/personas/>

Integrated cataloguing

- Callie has a database server that she uses to analyze relational data.
- She creates a new sandbox for each type of analysis.
- However, she often forgets to catalog her sandboxes.

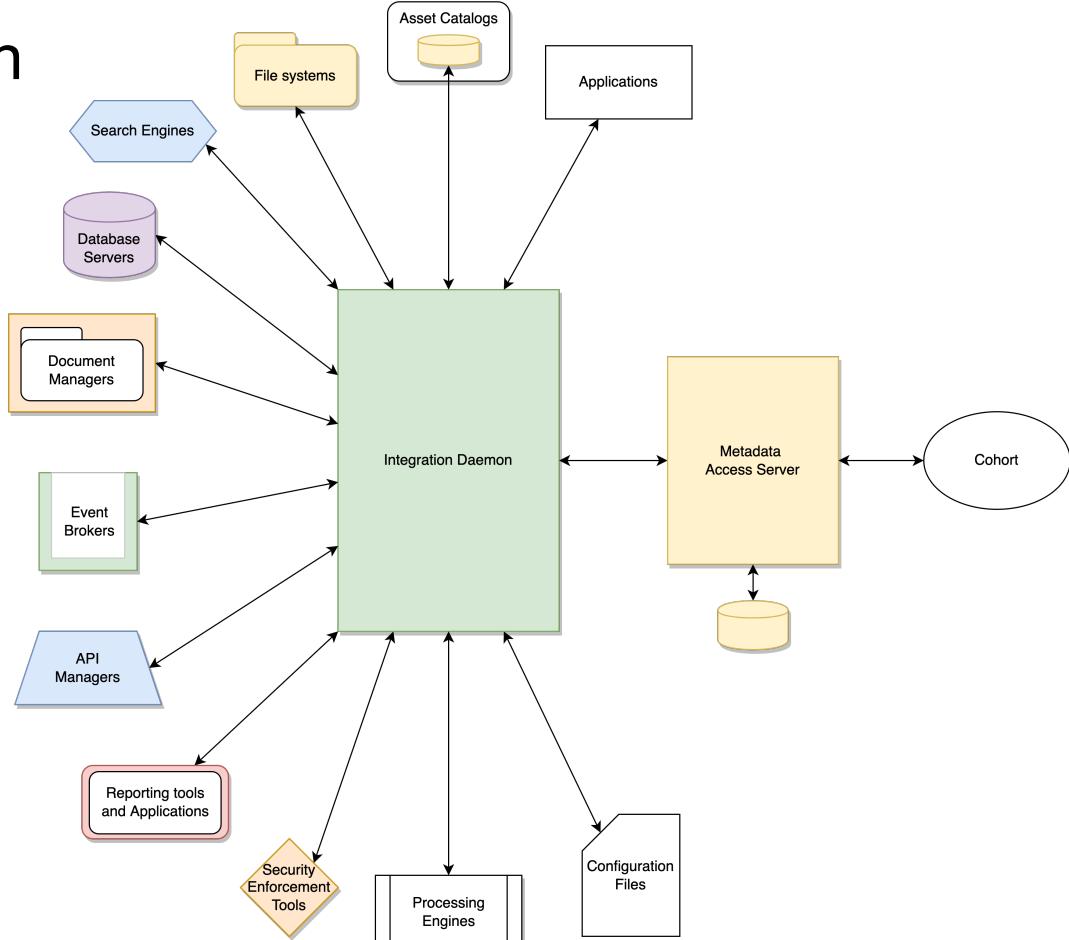


Automated cataloguing of Callie's Sandboxes

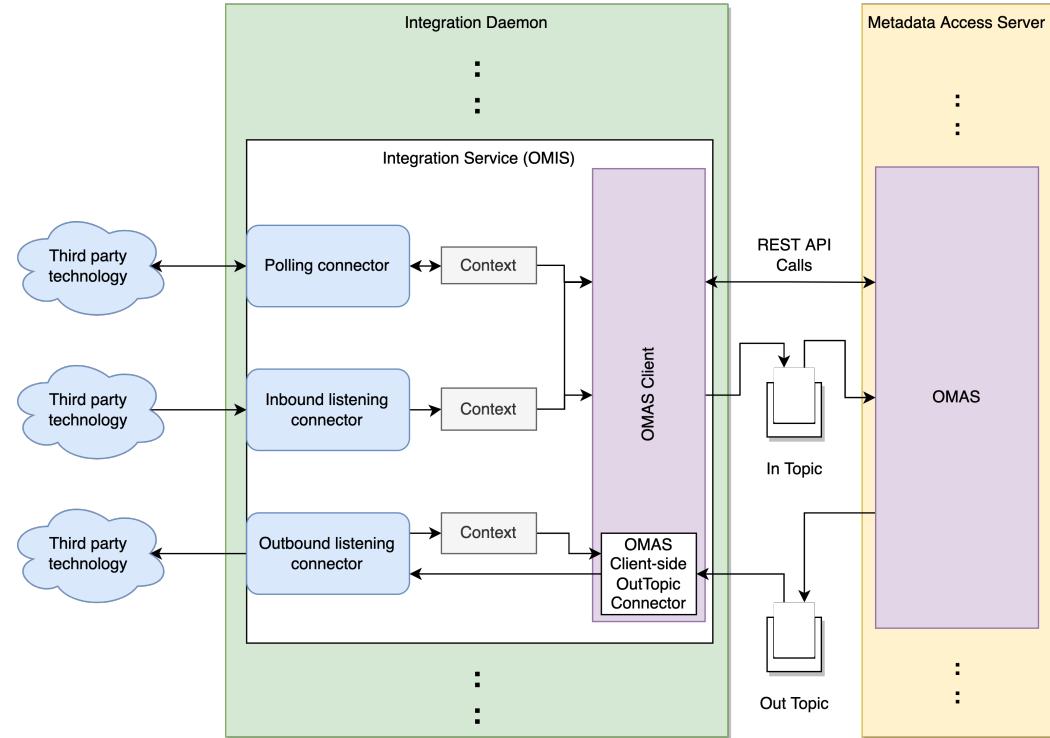


The Integration Daemon

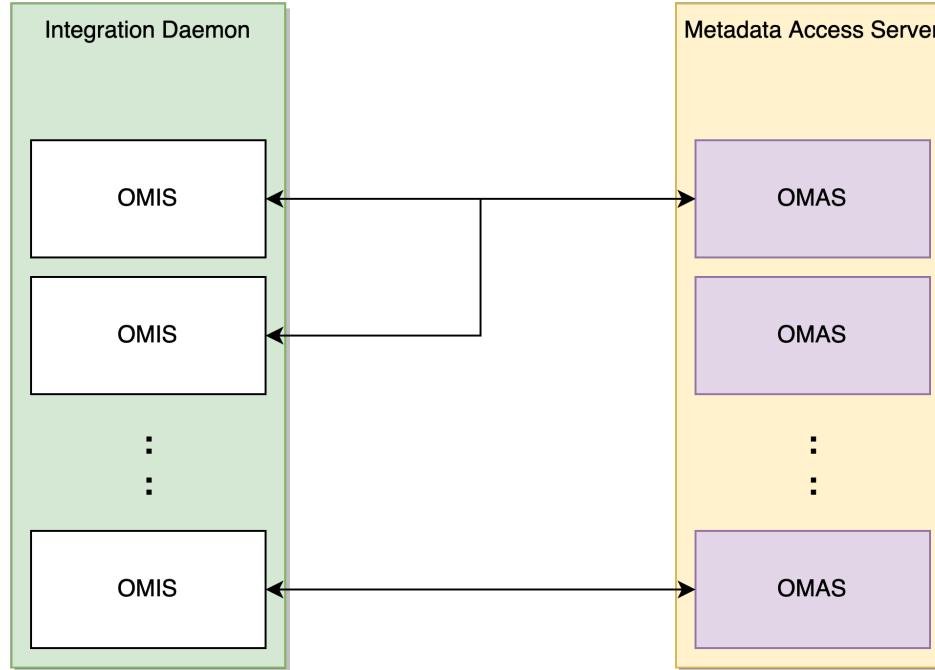
- A type of OMAG Server
- Metadata extraction, capture and delivery



Inside the integration daemon

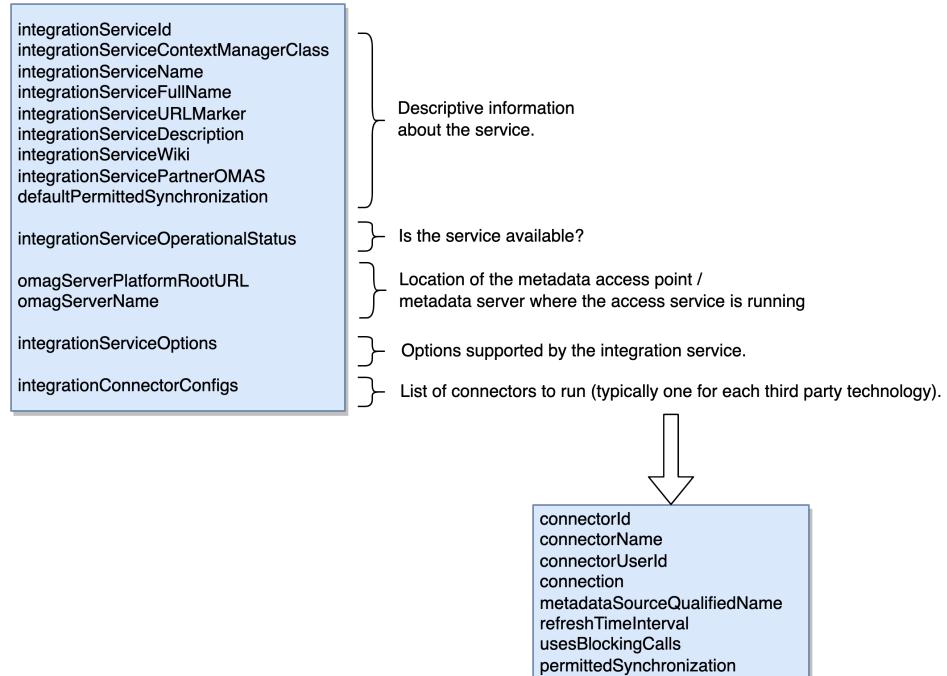


Supporting the metadata needs of different technologies

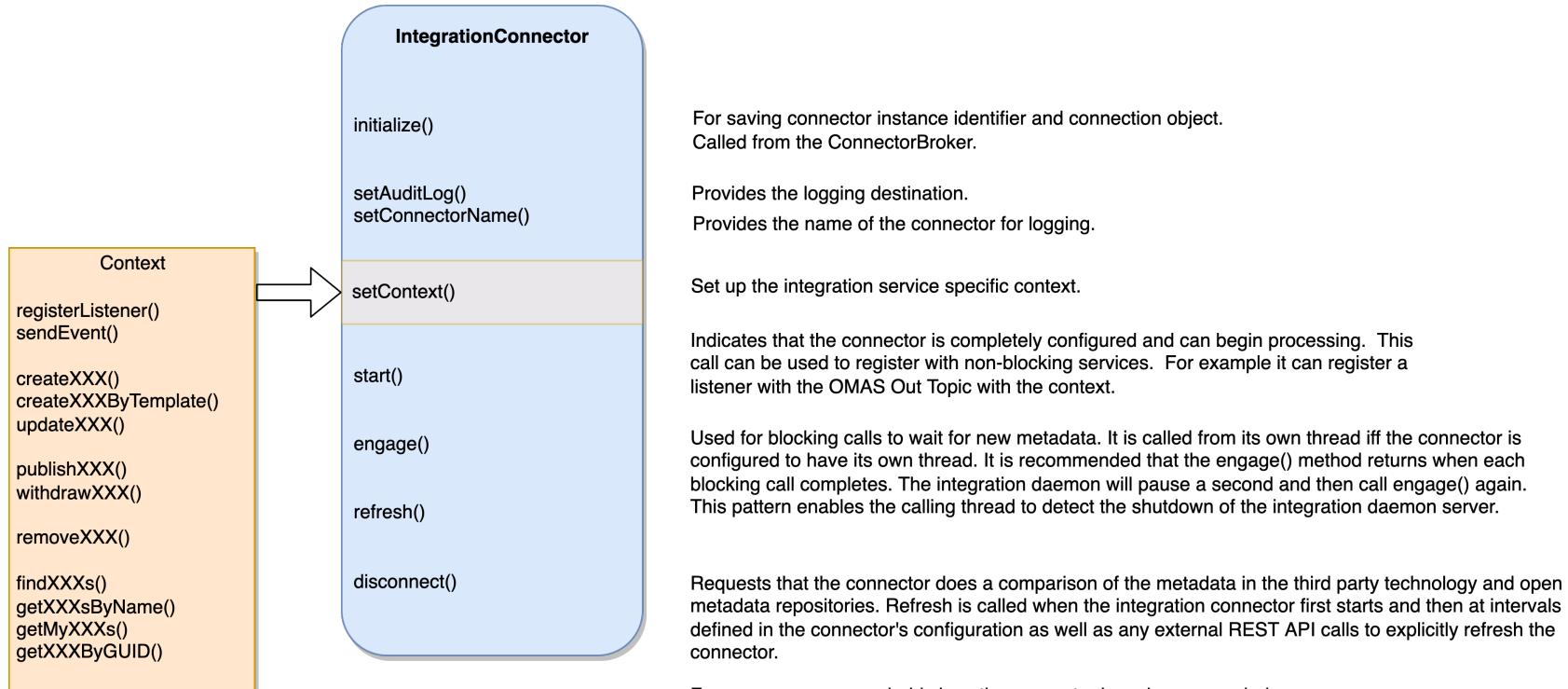


Integration Connector configuration

- The configuration provides the integration daemon with the information it needs to control the lifecycle and runtime support needed by the connector.



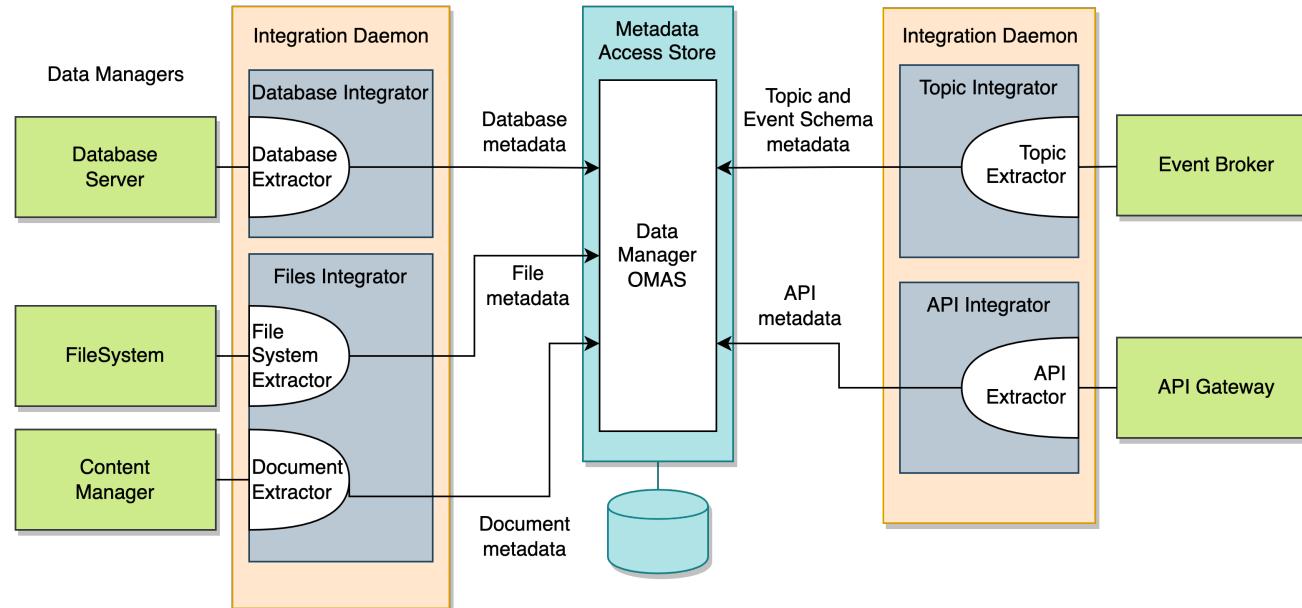
Integration Connector Implementation



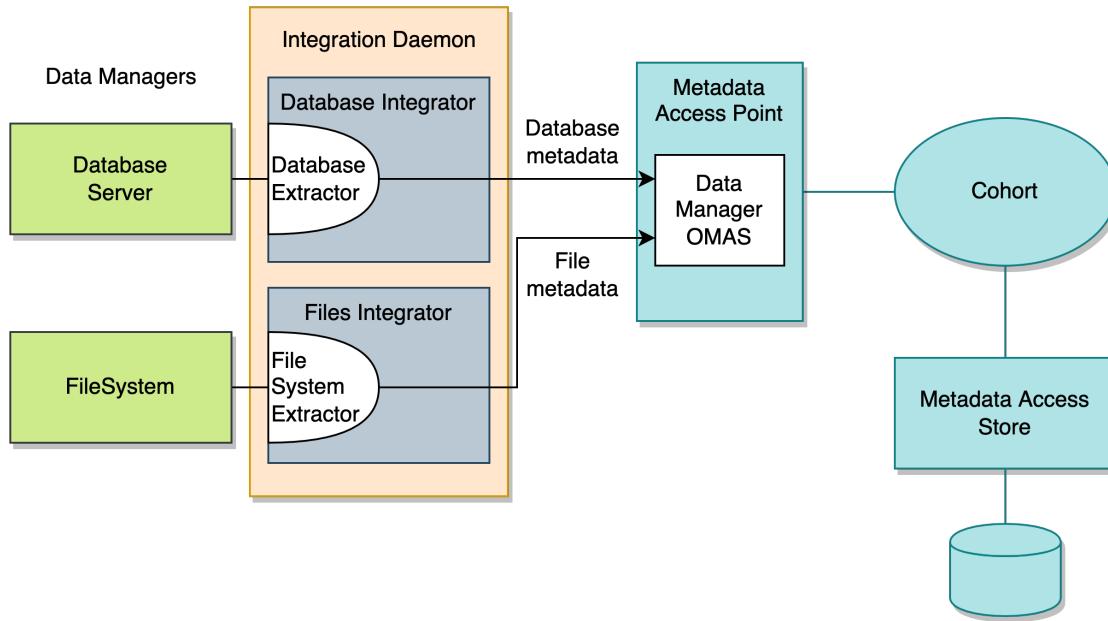
Metadata Extraction

Legend:

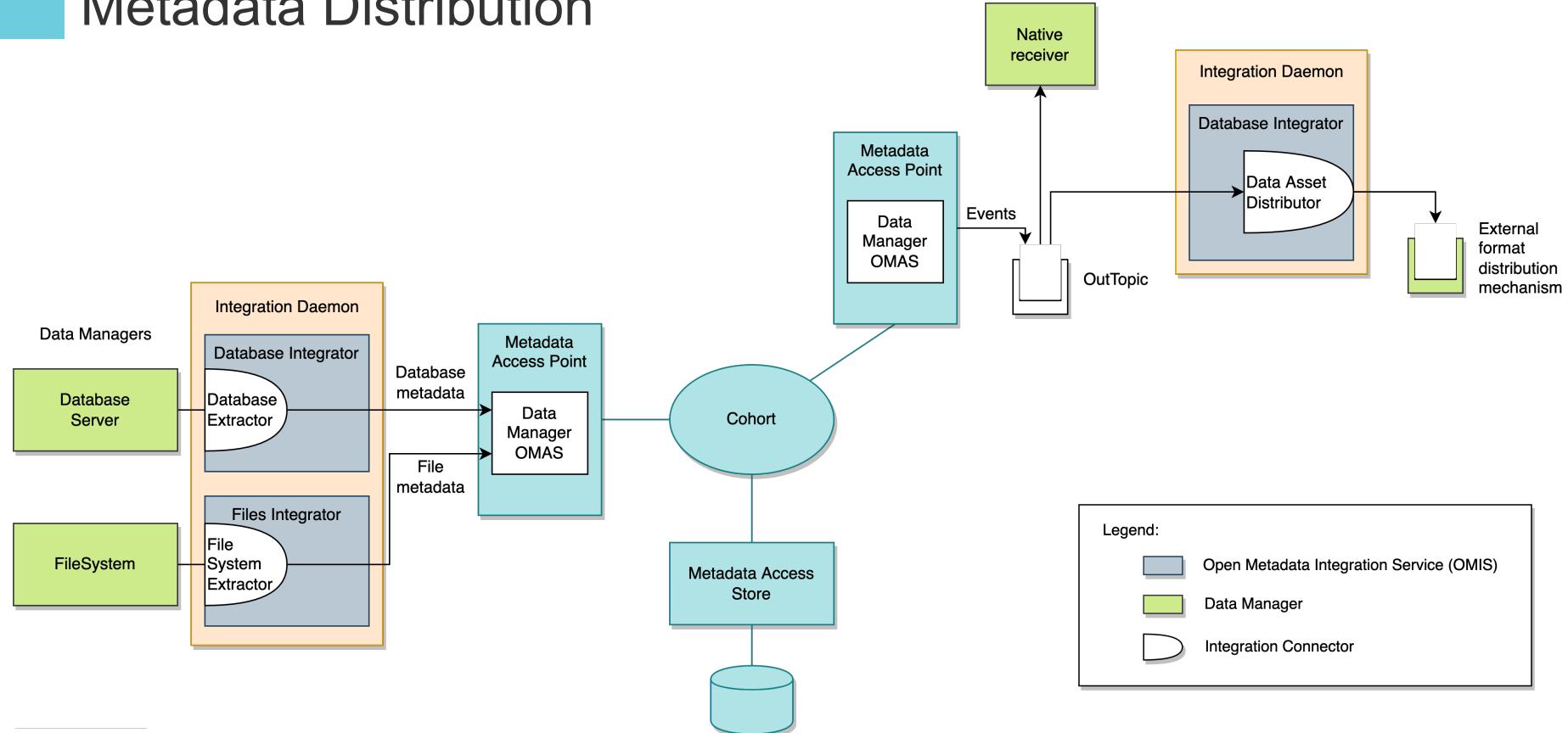
- Open Metadata Integration Service (OMIS)
- Data Manager
- Integration Connector



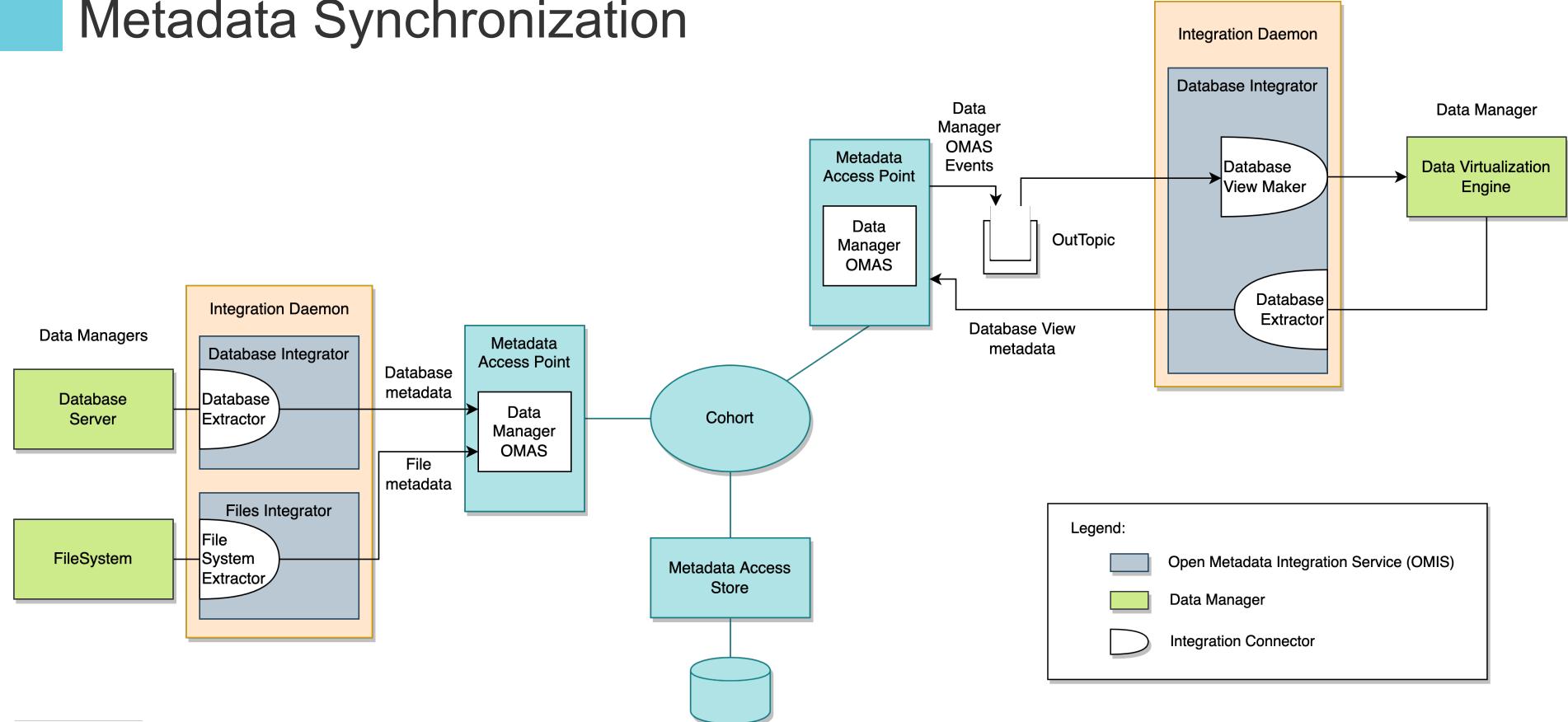
Metadata Extraction (remote store)



Metadata Distribution



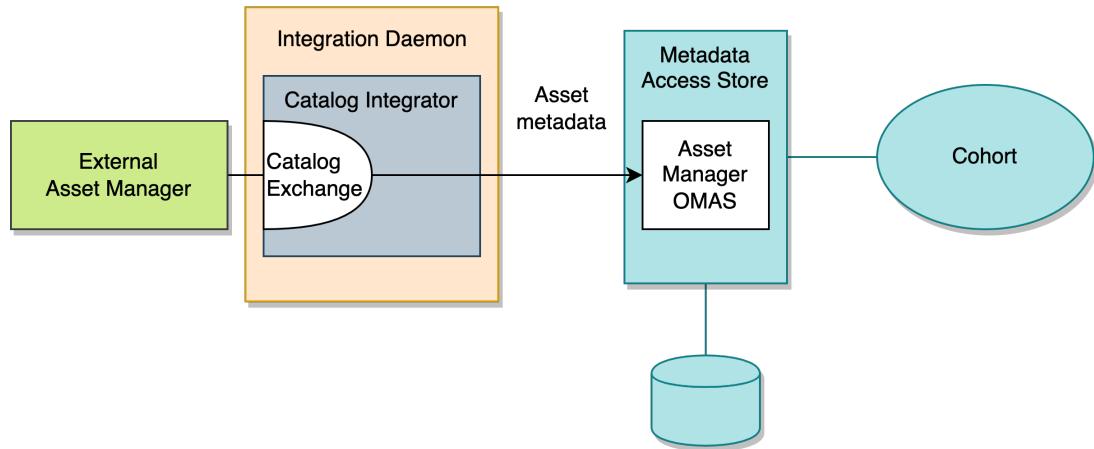
Metadata Synchronization



External Asset Manager

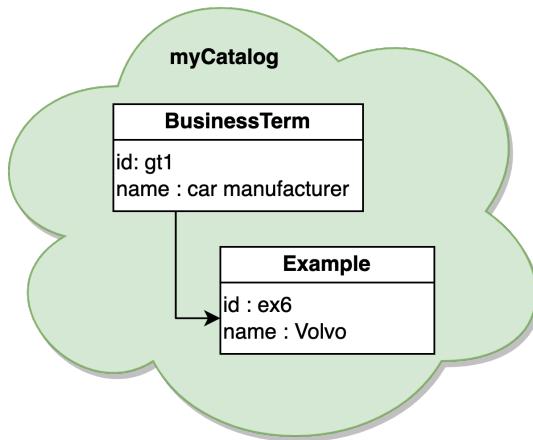
Legend:

- Open Metadata Integration Service (OMIS)
- External Asset Manager
- Integration Connector

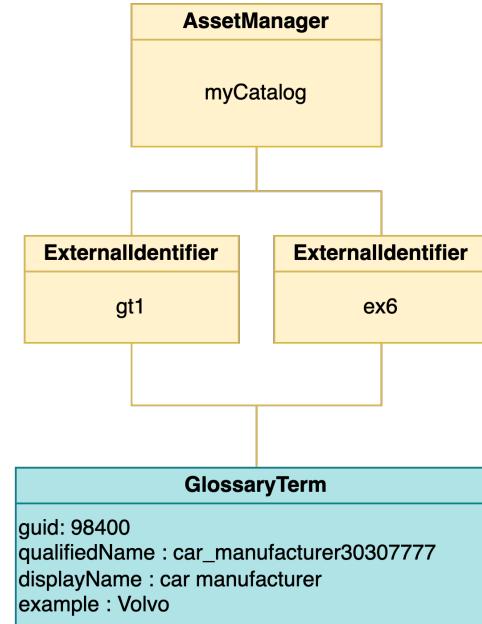


Many-to-One Mapping

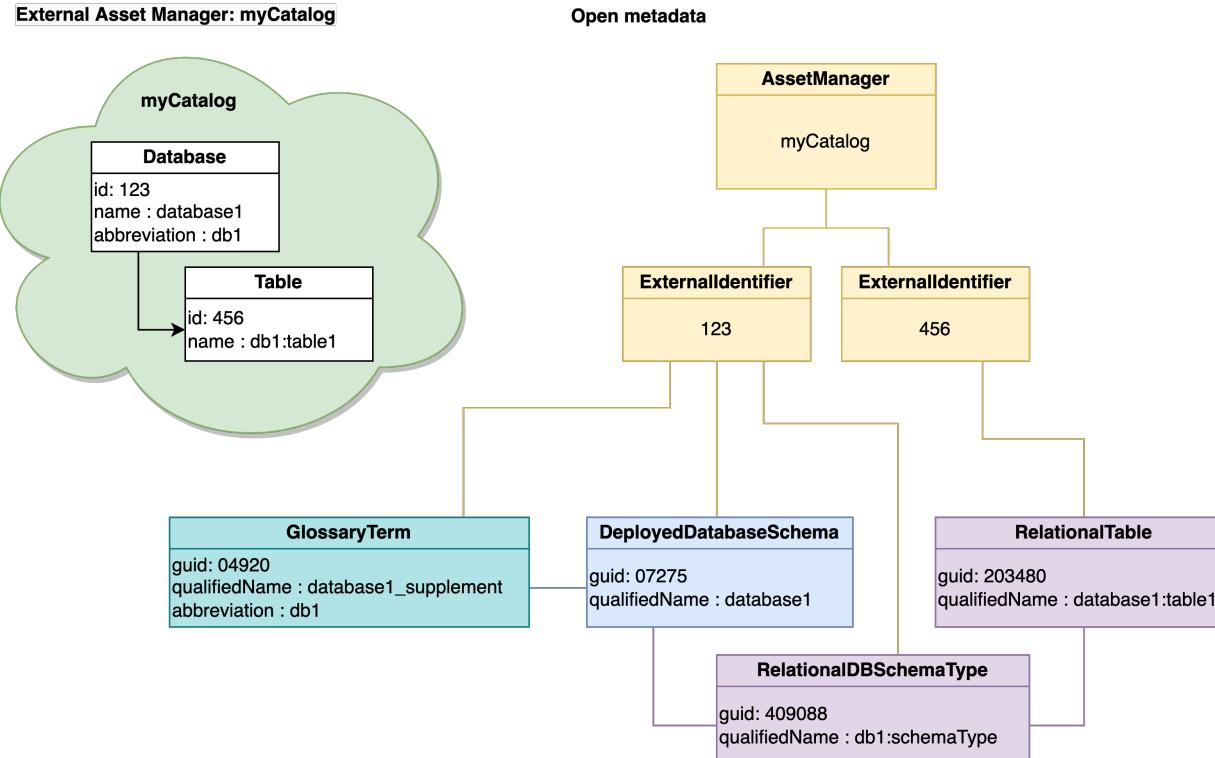
External Asset Manager: myCatalog



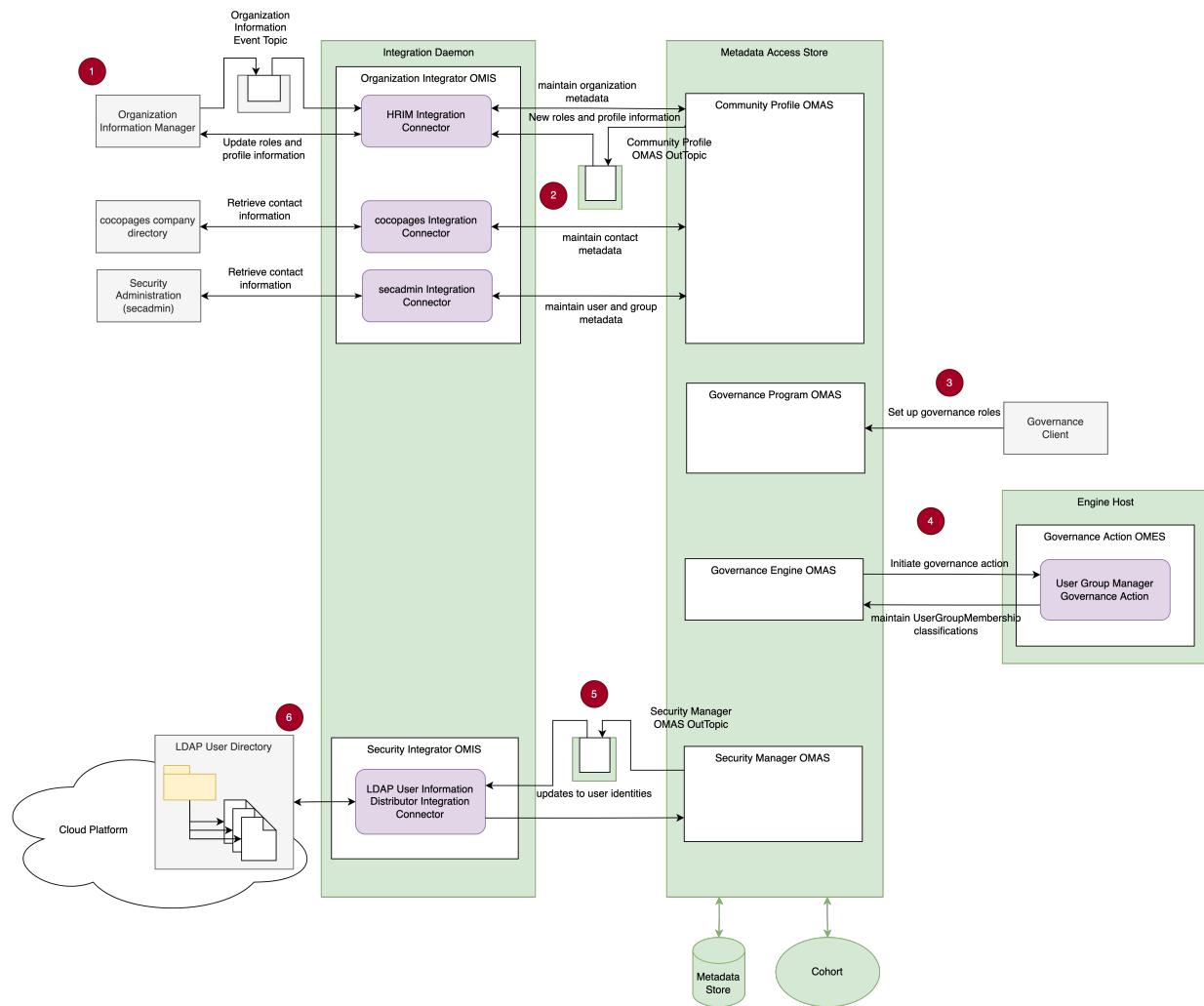
Open metadata



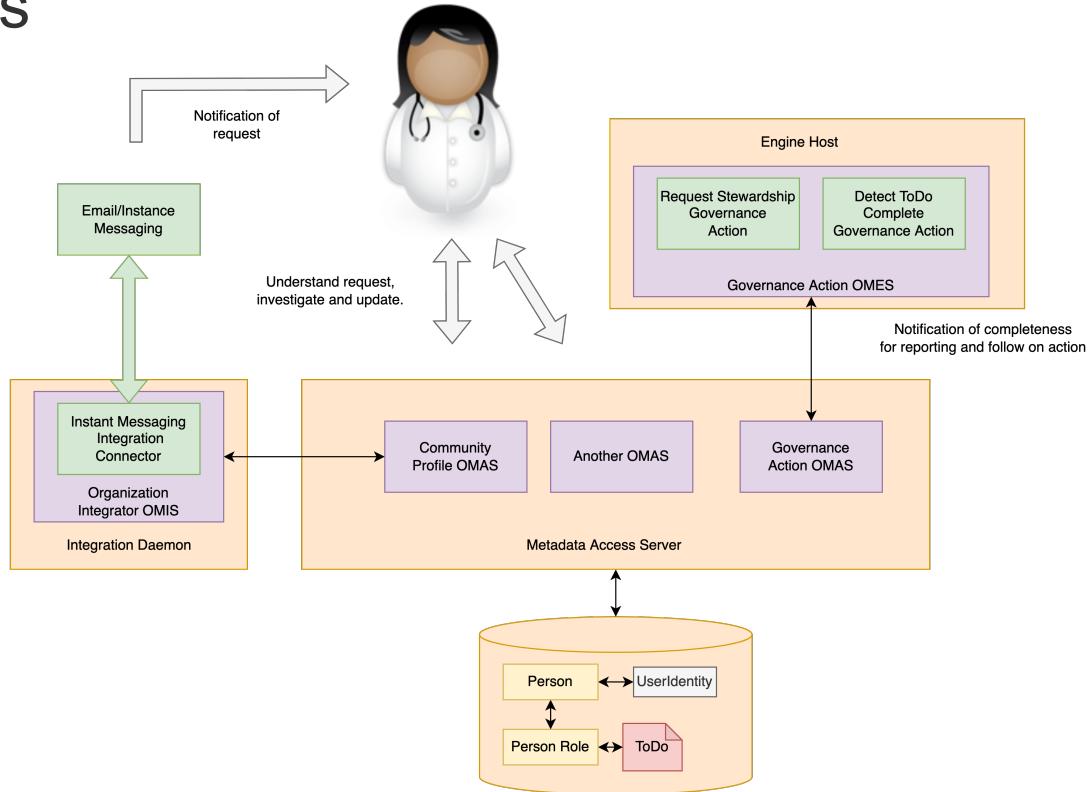
One-to-Many Mapping



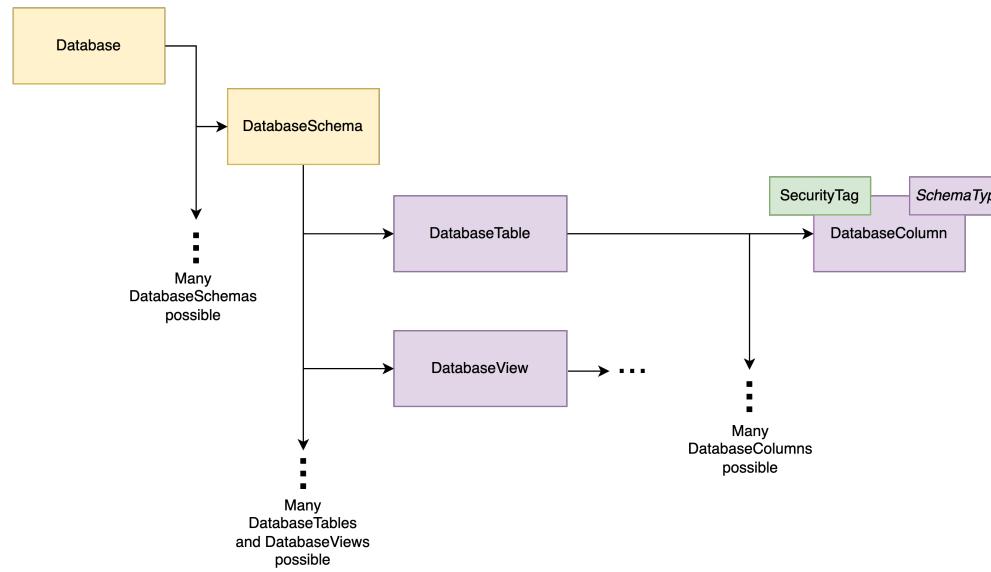
Onboarding an organization

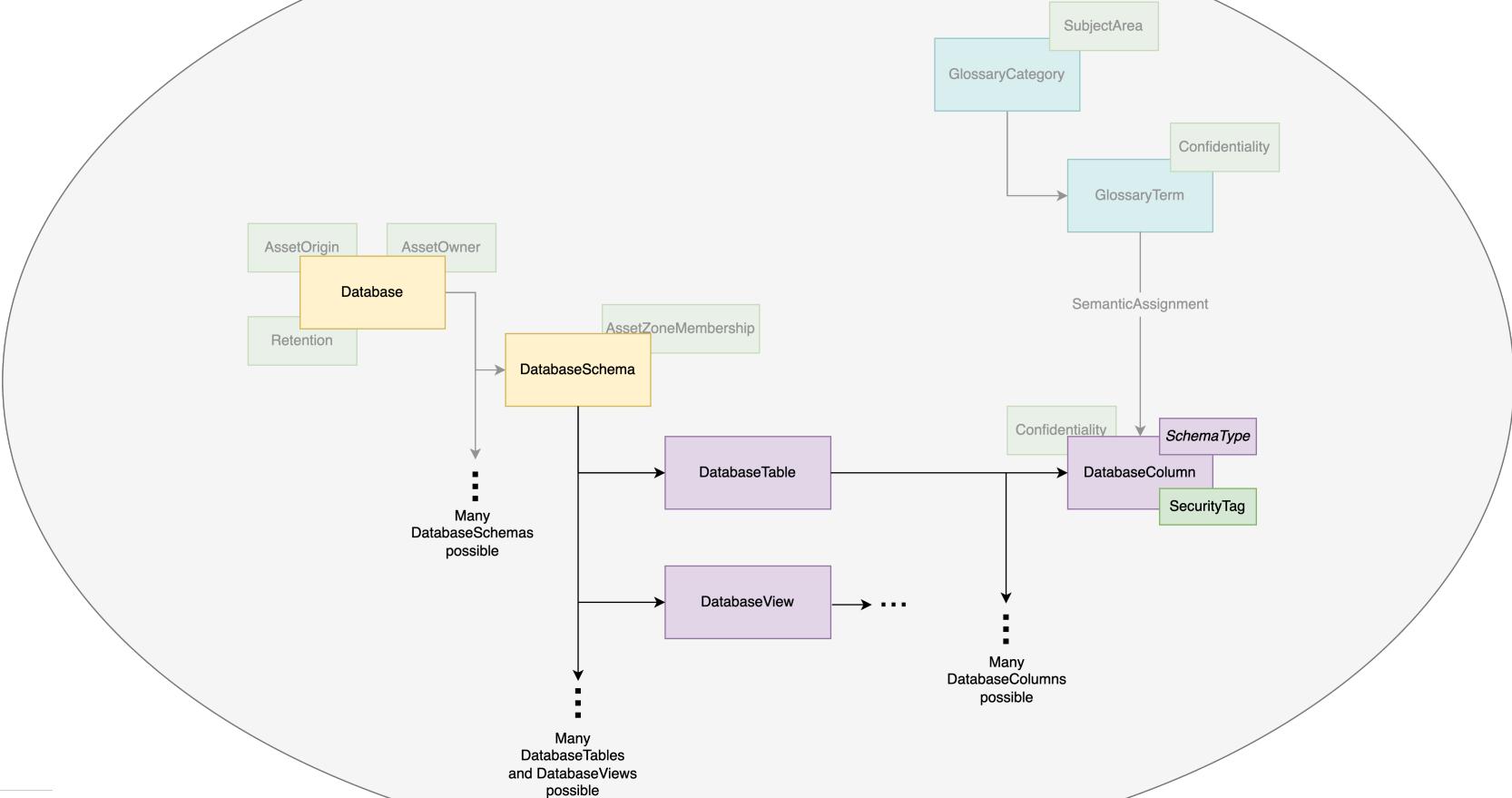


Steward Notifications

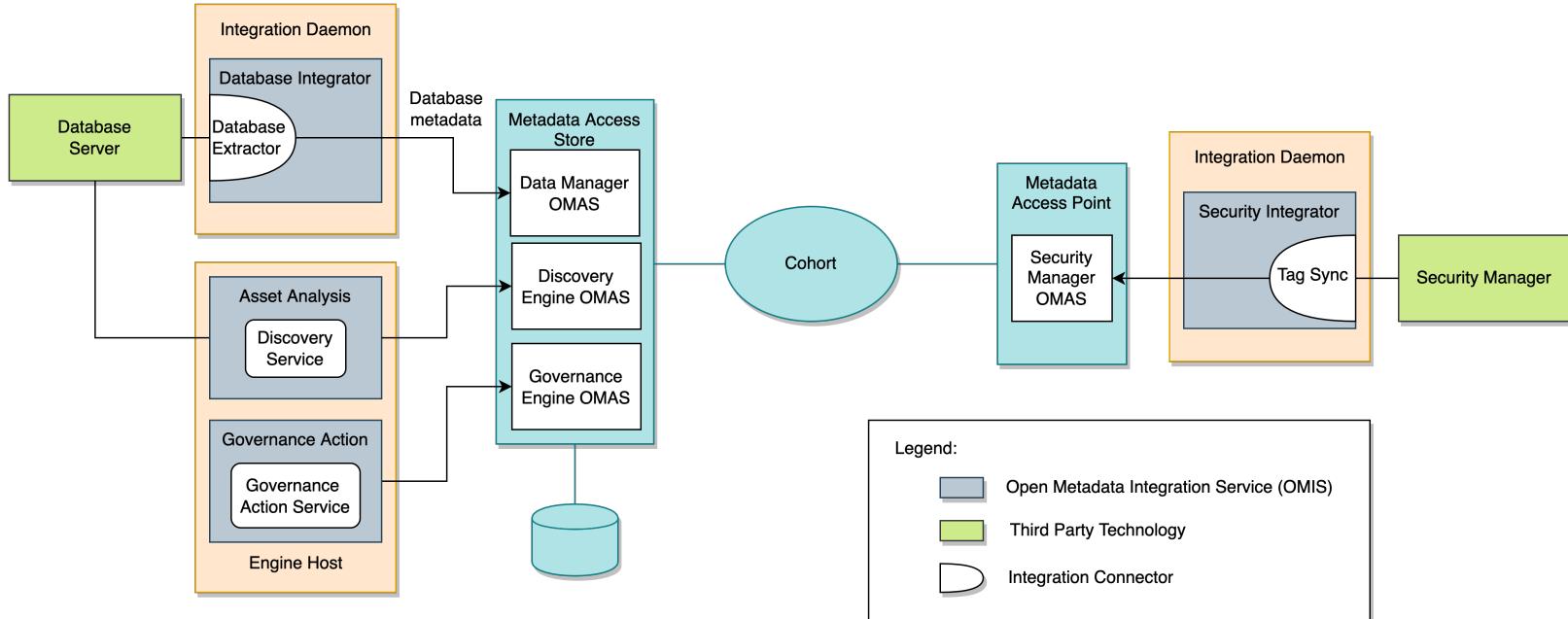


Pushing security tags to security managers

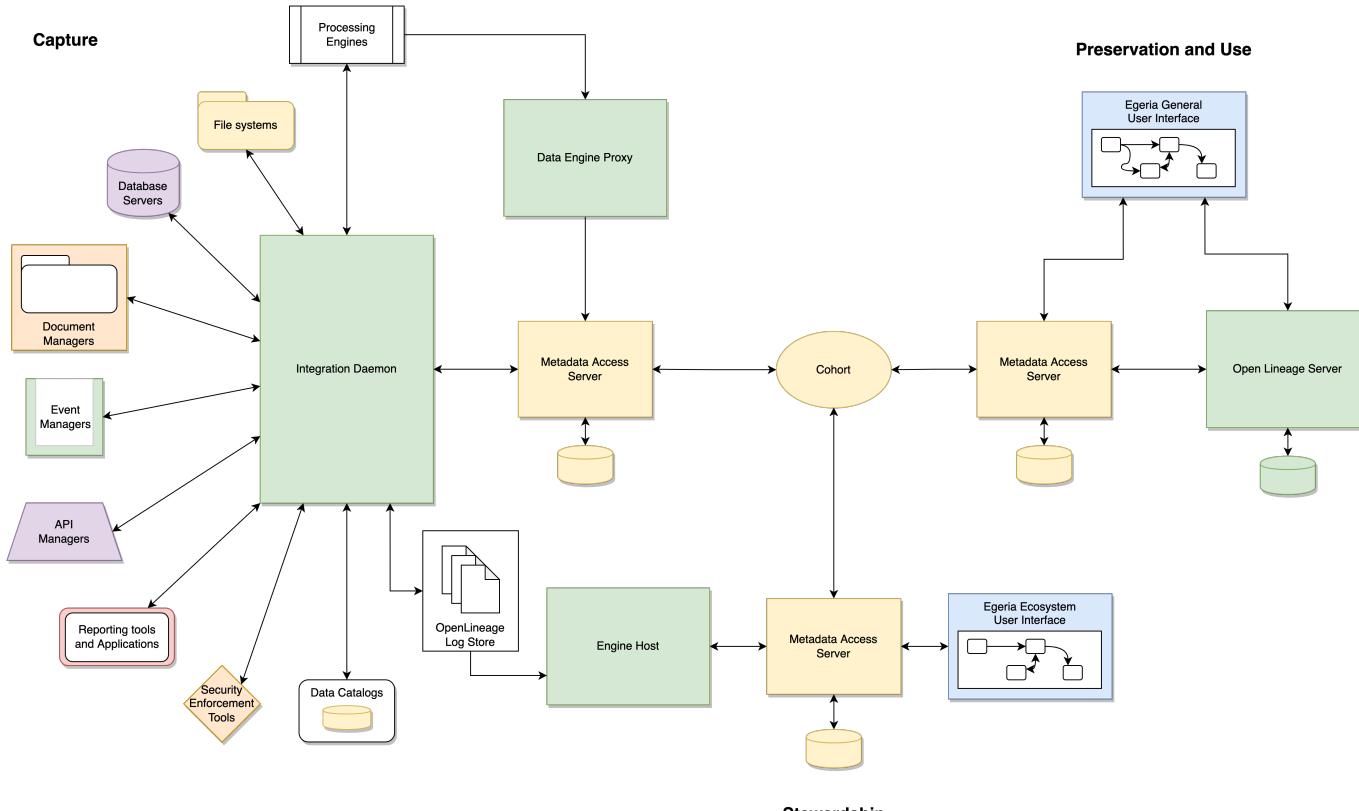




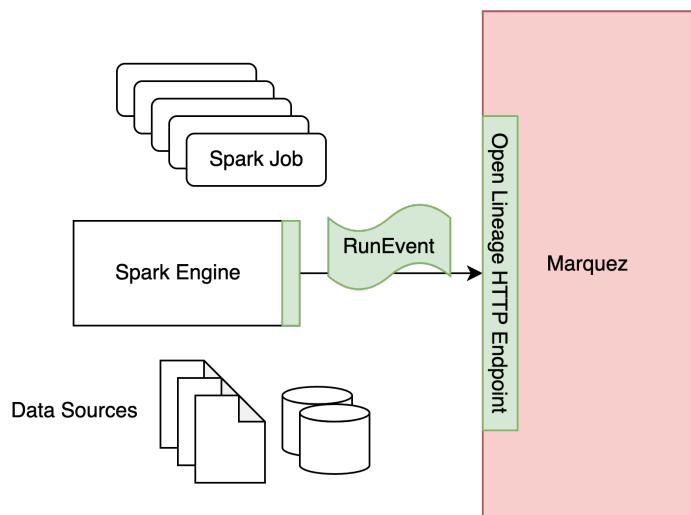
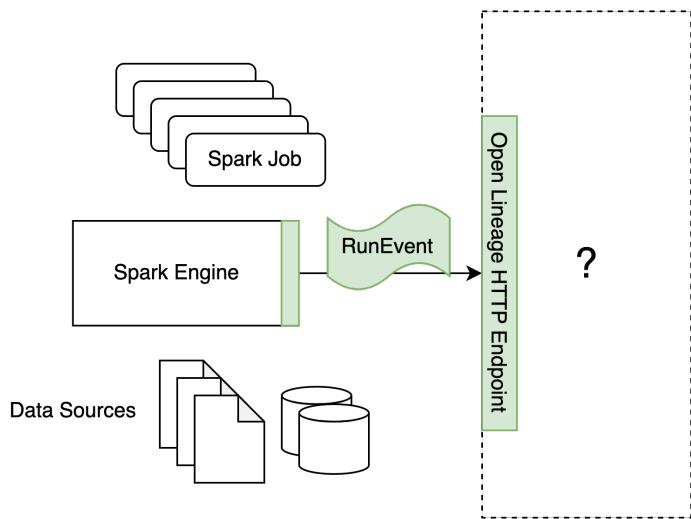
Pushing security tags to a security manager



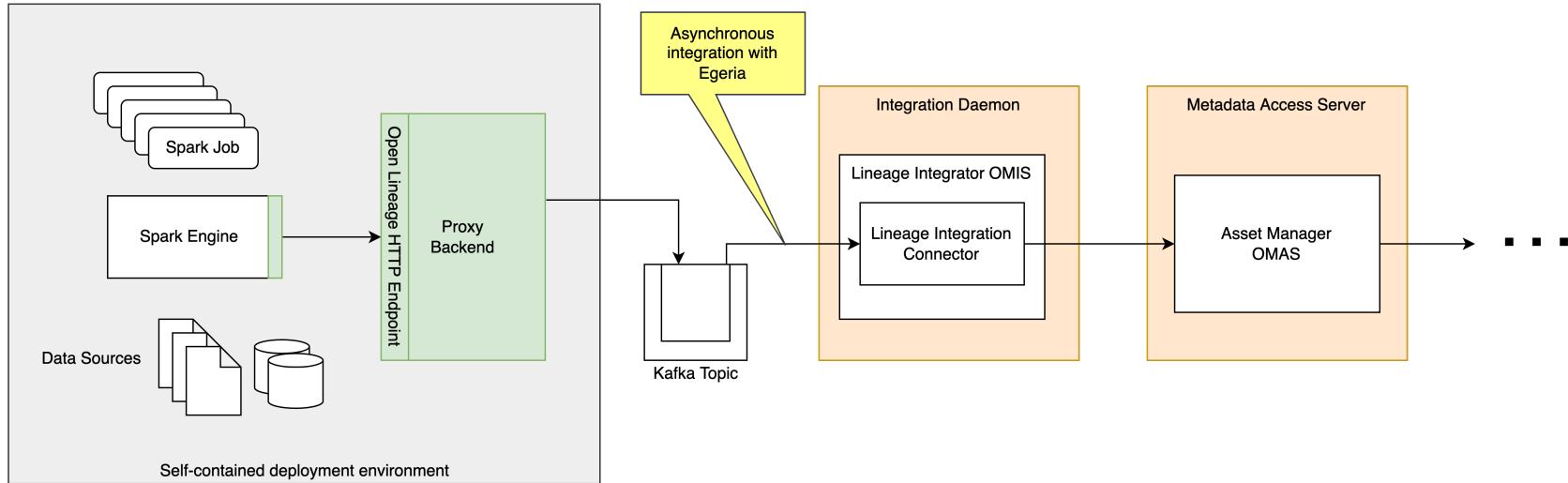
Lineage management



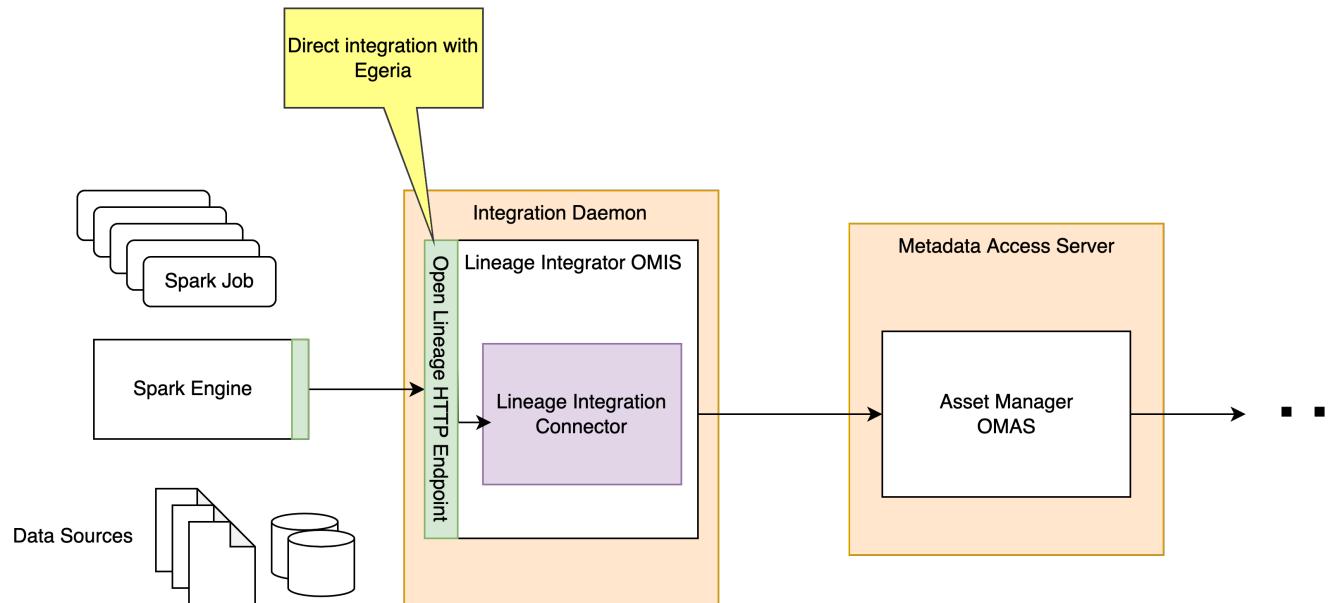
OpenLineage standard



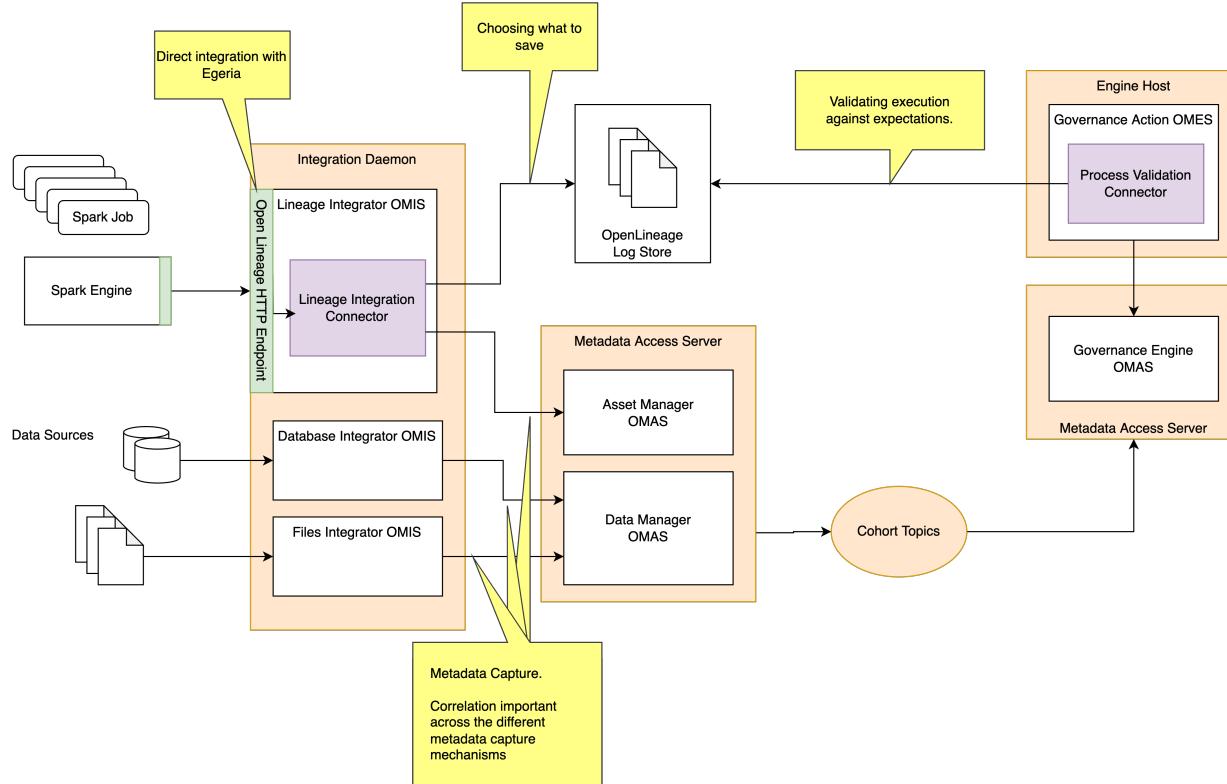
Receiving OpenLineage events



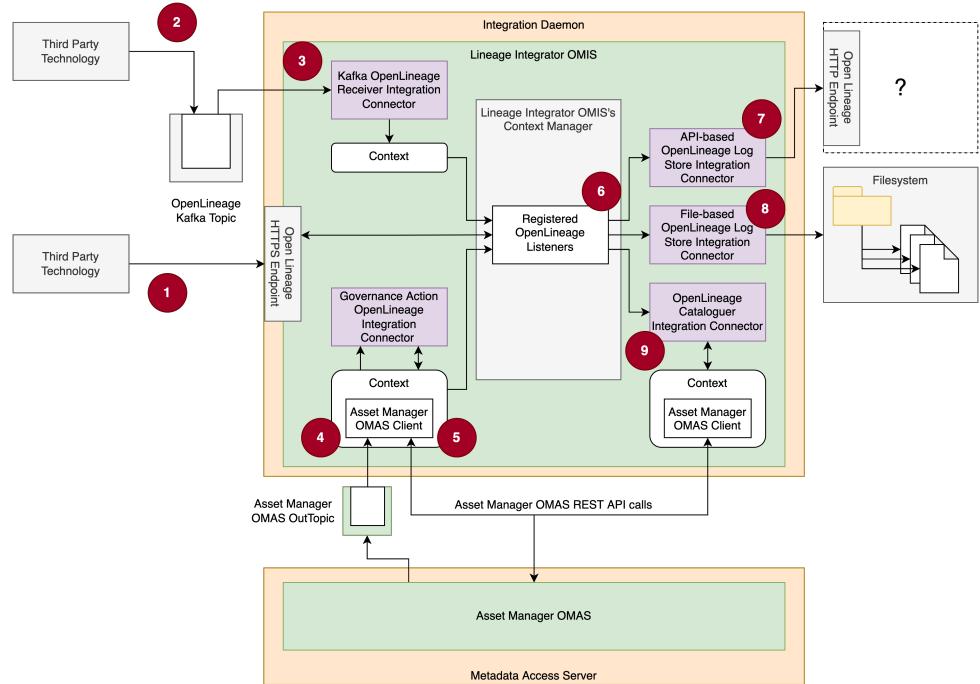
Direct receipt of OpenLineage events



Lineage Deployment



Inside the Lineage Integrator OMIS



Date	time	Title	Description	Presenter
January 2022	n/a	n/a	There is no webinar this month. The team is working on creating a new set of monthly Dojo sessions . The first of these sessions will be on the 17th of January 2022 10AM - 17:00PM (UTC) and will be on How to setup and run Egeria in different environments . This will include the Kubernetes environment. Zoom Conference https://zoom.us/j/523629111	
7th February 2022	15:00 UTC	Using an integration connector	This session covers how to use Integration connectors to connect technologies into Egeria..	Mandy Chessell
7th March 2022	15:00 UTC	How to build an integration connector	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 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.	Mandy Chessell
4th April 2022	15:00 UTC	Using a repository connector	This session covers how to use Repository Connectors to connect technologies into Egeria; focussing on XTDB (formerly known as crux). Every wanted to know what the state of your metadata was at some specific time in the past? This session will introduce the XTDB open metadata repository that supports these historical metadata queries.	Chris Grote

Open forum



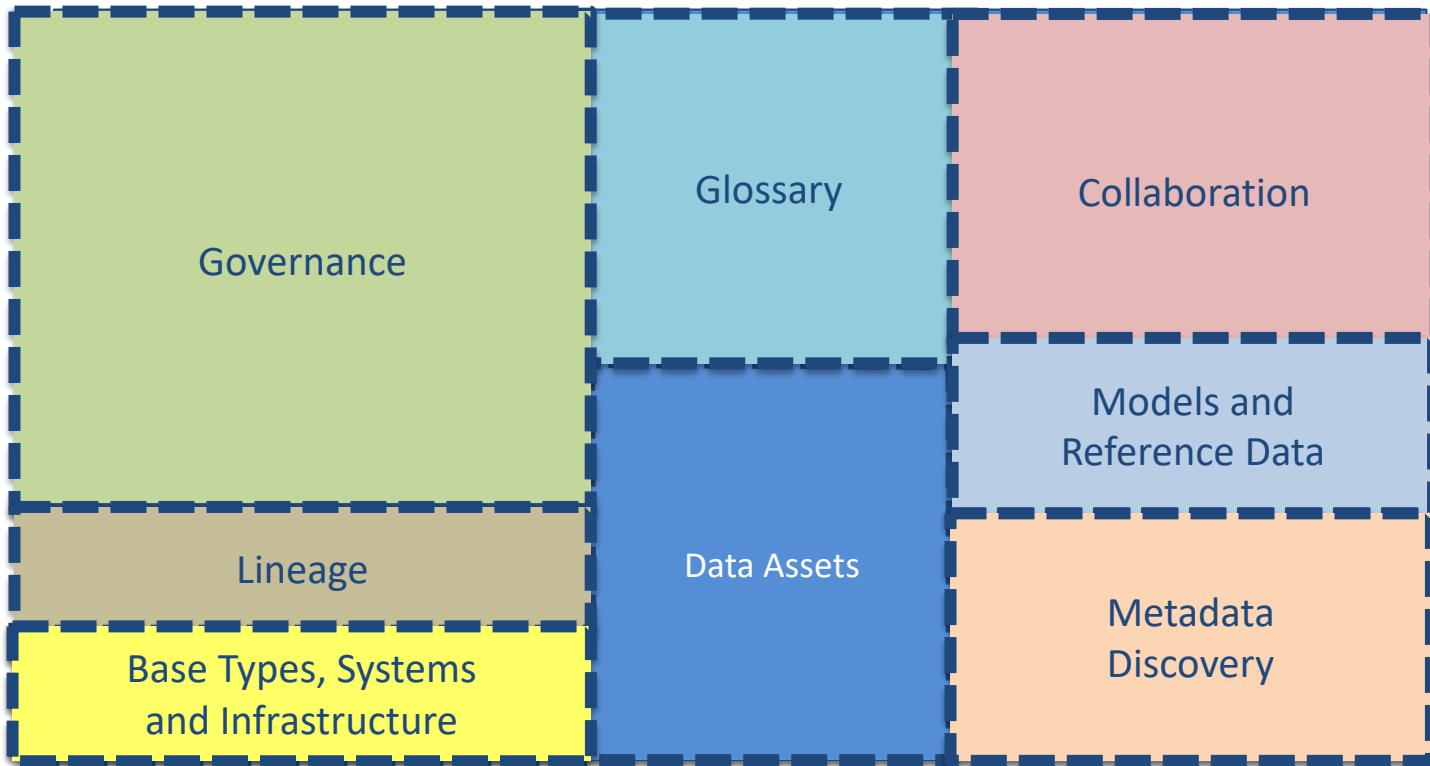
THANK YOU!



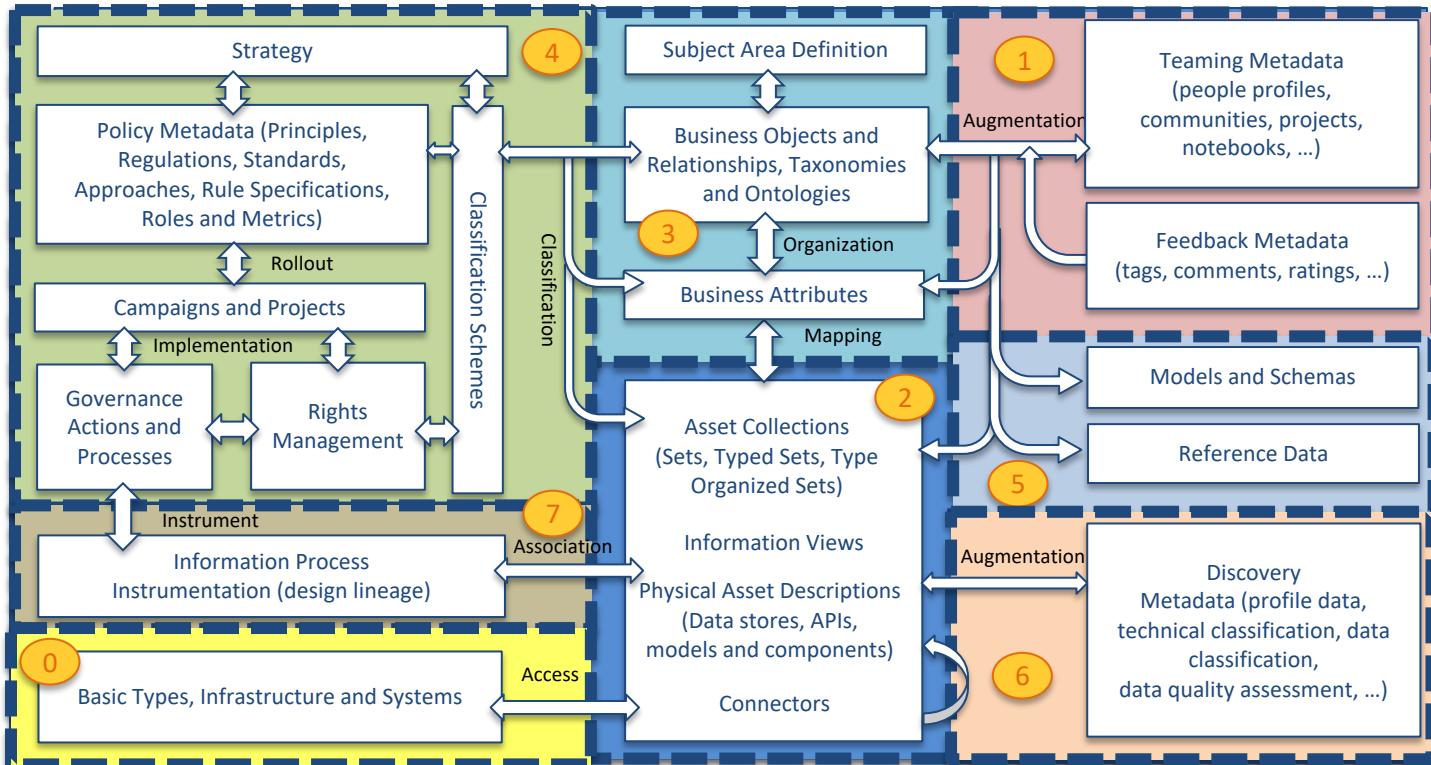
Achievements

- 700 linked open metadata types demonstrating how the knowledge from many tools can be linked together.
- Open metadata repository interface proven for table, graph and hierarchical DB stores.
- Enterprise queries and replication across heterogeneous technologies
- Conformance test suite and mark
- Automated configuration of data virtualization technology and security as new data sets are added to a data lake
- Suite of persona-based labs and tutorial using Jupyter Notebooks.
- Virtual graph of metadata maintained across distributed heterogenous metadata repositories.
- Frameworks, APIs and connectors for minimizing integration cost for different types of technologies
- Virtual repository explorer UI
- Instance based security
- Controlling visibility of assets through zones
- Scalable, secure platform configurable and customizable through connectors
- Purpose-based data access
- Metadata versioning and provenance
- Multi-tenant UI based on carbon
- W3C semantic standards pattern for data model exchange
- Automation of metadata acquisition through templates, daemons, discovery services and stewardship.
- Classification of assets
- Reference data management
- Multi-technology collaboration and feedback
- Multi-domain governance model
- Digital service lifecycle, from business design, development, devOps and use.
- Comprehensive open lineage services.
- Metadata deduplication

Scope of metadata covered



Scope of metadata covered



<https://egeria.odpi.org/open-metadata-publication/website/open-metadata-types/>

<https://egeria-project.org>

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.

