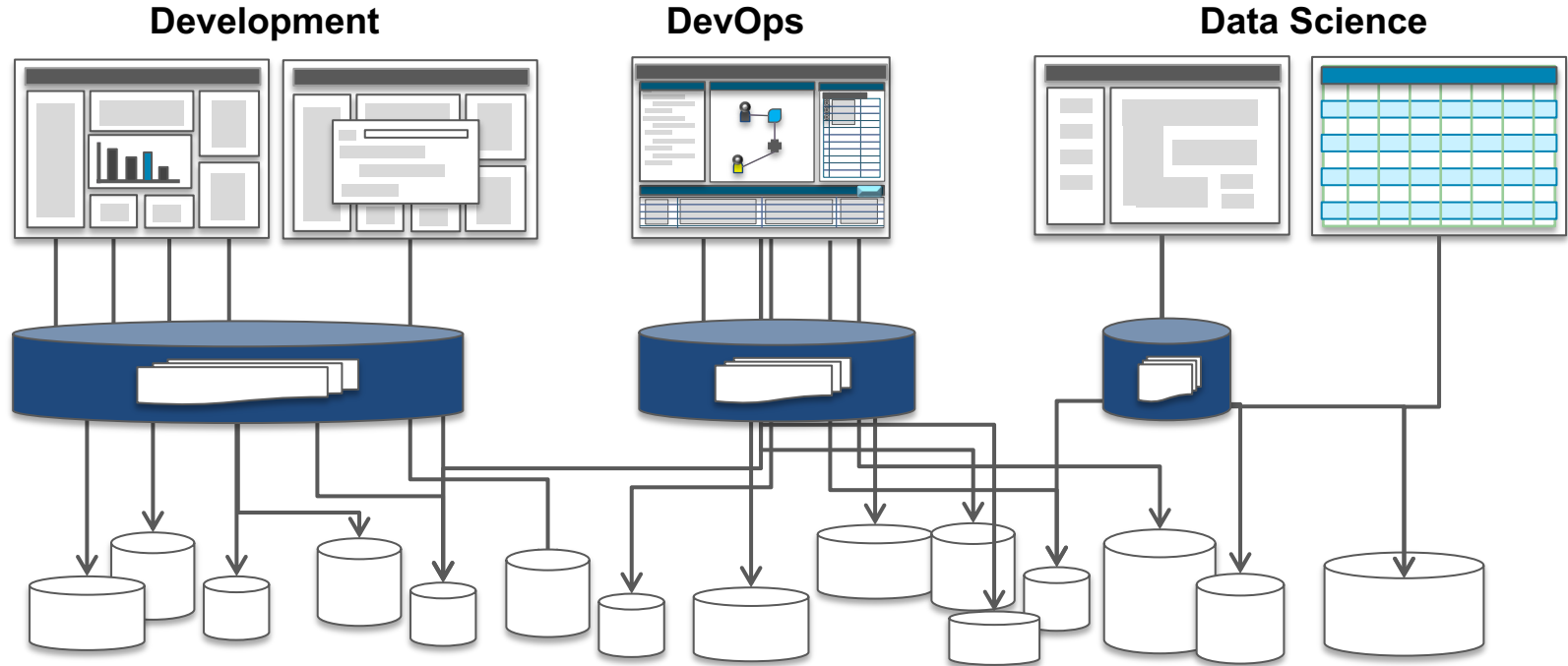


For Dojo Day 1

EGERIA AND OPEN METADATA

Mandy Chessell CBE FREng
Egeria Open Source Project Lead

Today's reality – organizations buy lots of tools



Egeria's value



- Egeria enables the sharing of knowledge between people and processes that use different tools and technologies
 - Increasing agility, collaboration and the availability of relevant data when making decisions.
 - Breaking down silos between traditional technologies such as data and applications; applications and infrastructure management; security and data.

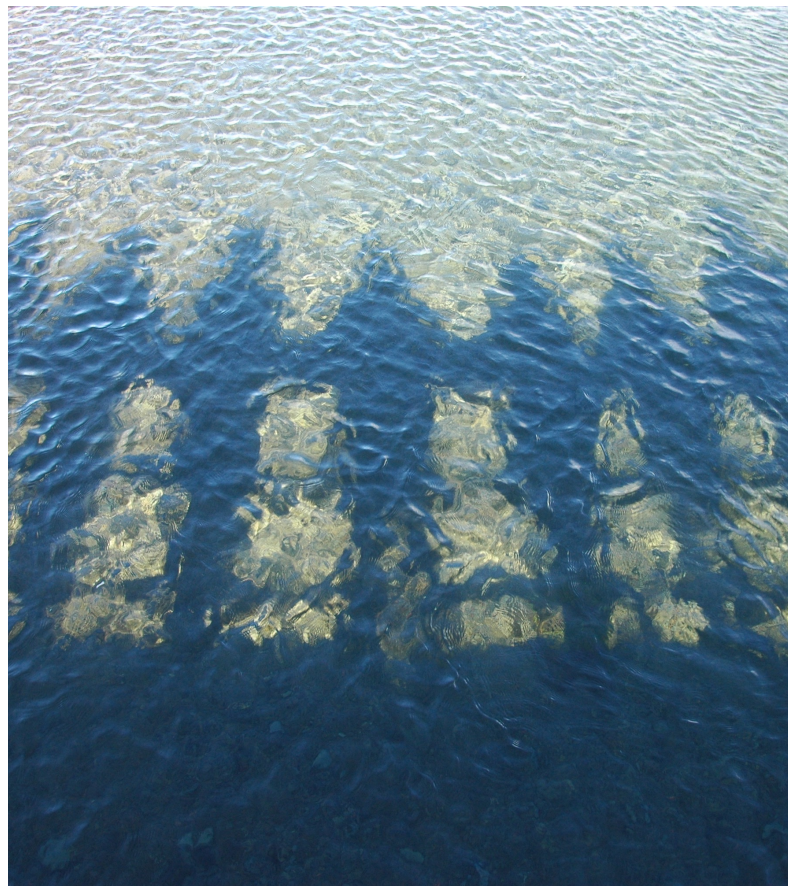
This is not a new problem ...

- Over the years, our industry has gained plenty of experience discovering what does not work
 - Centralized metadata repository
 - “Just use our tools” – Single vendor lock in
 - Paper standards that each vendor implements
 - Bridges that copy metadata between tools
 - Bespoke point-to-point integrations between tools

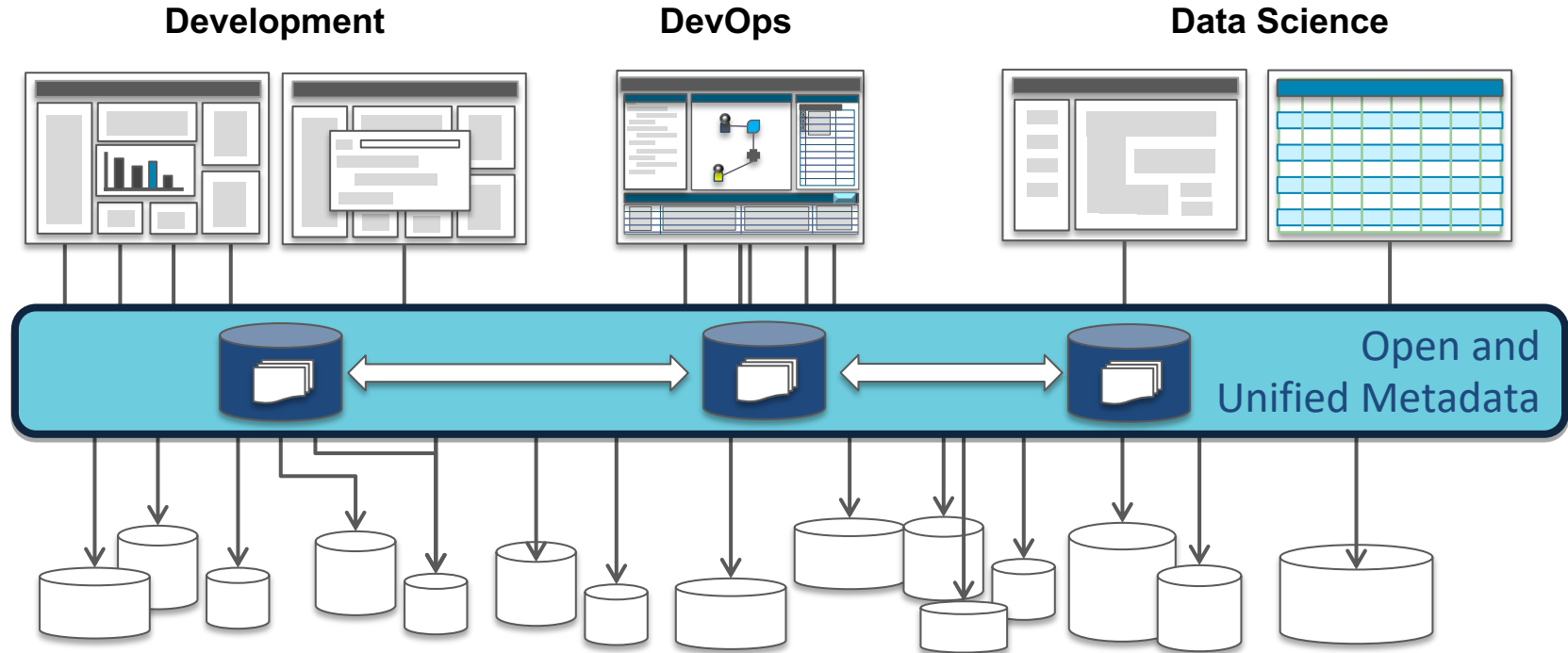


What is different about Egeria

- Open source, multi-vendor and client collaboration
- Peer-to-peer architecture; retaining the value of each technology
- Self-configuring, distributed, real-time exchange of knowledge
- Scales from Raspberry Pi to multi-instance HA cloud deployment
- Multi-tenant
- Instance based security
- Customizable through connectors



Egeria enables exchange of metadata between tools from different vendors



What makes metadata integration hard?

- Metadata standards are many and each covers a small subset of the integration space
- Impedance mismatch in:
 - Terminology
 - Granularity
 - Capability
 - Availability
 - Technology
- Low priority in offering managers eyes → leads to lack of investment resulting in minimal capability in vendor offerings

Three layers of challenge for Egeria

Maintaining coherence of metadata across the ecosystem

Connectivity to enable the exchange of metadata

Infrastructure to host the open metadata and governance logic

Three layers of challenge for Egeria

Maintaining coherence of metadata across the ecosystem

Deduplication Profiling Classification Monitoring Provisioning Triage Validation Enforcement Remediation Enrichment

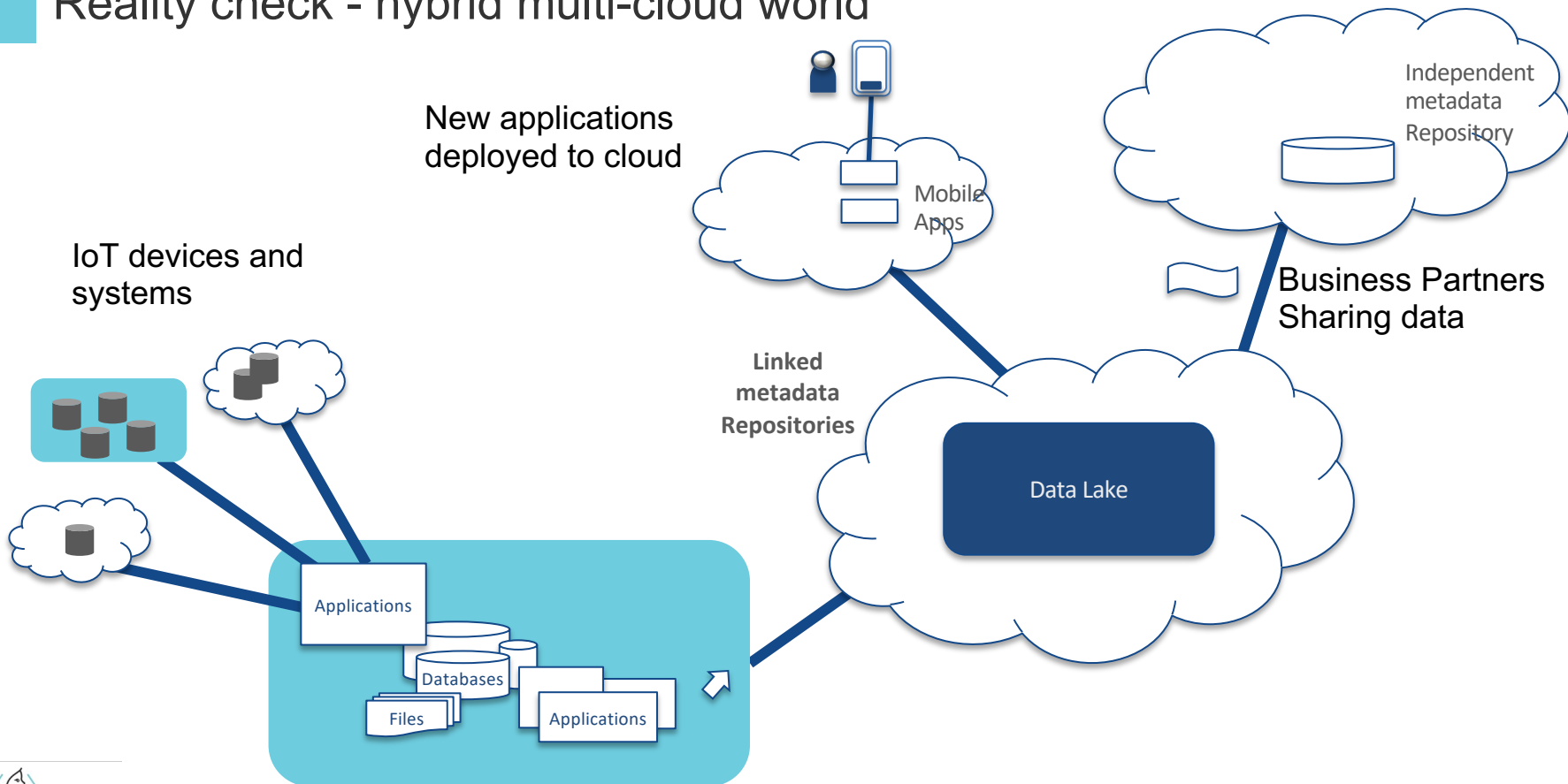
Connectivity to enable the exchange of metadata

Open Metadata Repository Cohort and Federated Queries Versioning and Provenance of Metadata Open Metadata Archives
Connector Interfaces and hosting Multi-layered security Specialist APIs and Topics Effectivity Dating User Interfaces
Metadata capture, transformation and storage Third party technology metadata exchange (poll, listen, push, pull) Type Definitions
Lineage and Mementos Curation and Templating Scheduling, Triggers and and Orchestration Audit Logging, Monitoring and Alerting

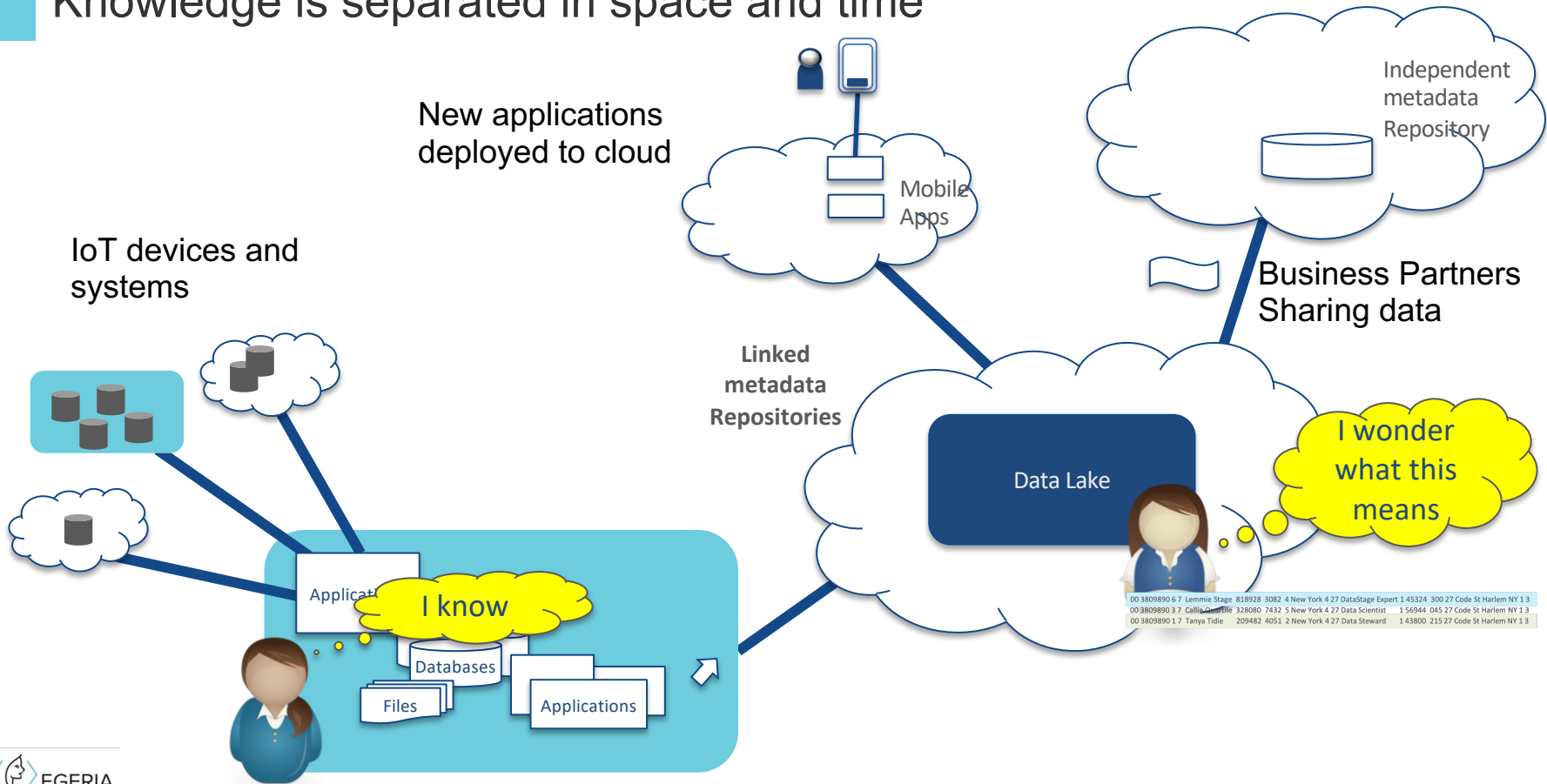
Infrastructure to host the open metadata and governance logic

Administration Multi-Tenancy REST Endpoint Server Management Containerization Transport-level Security

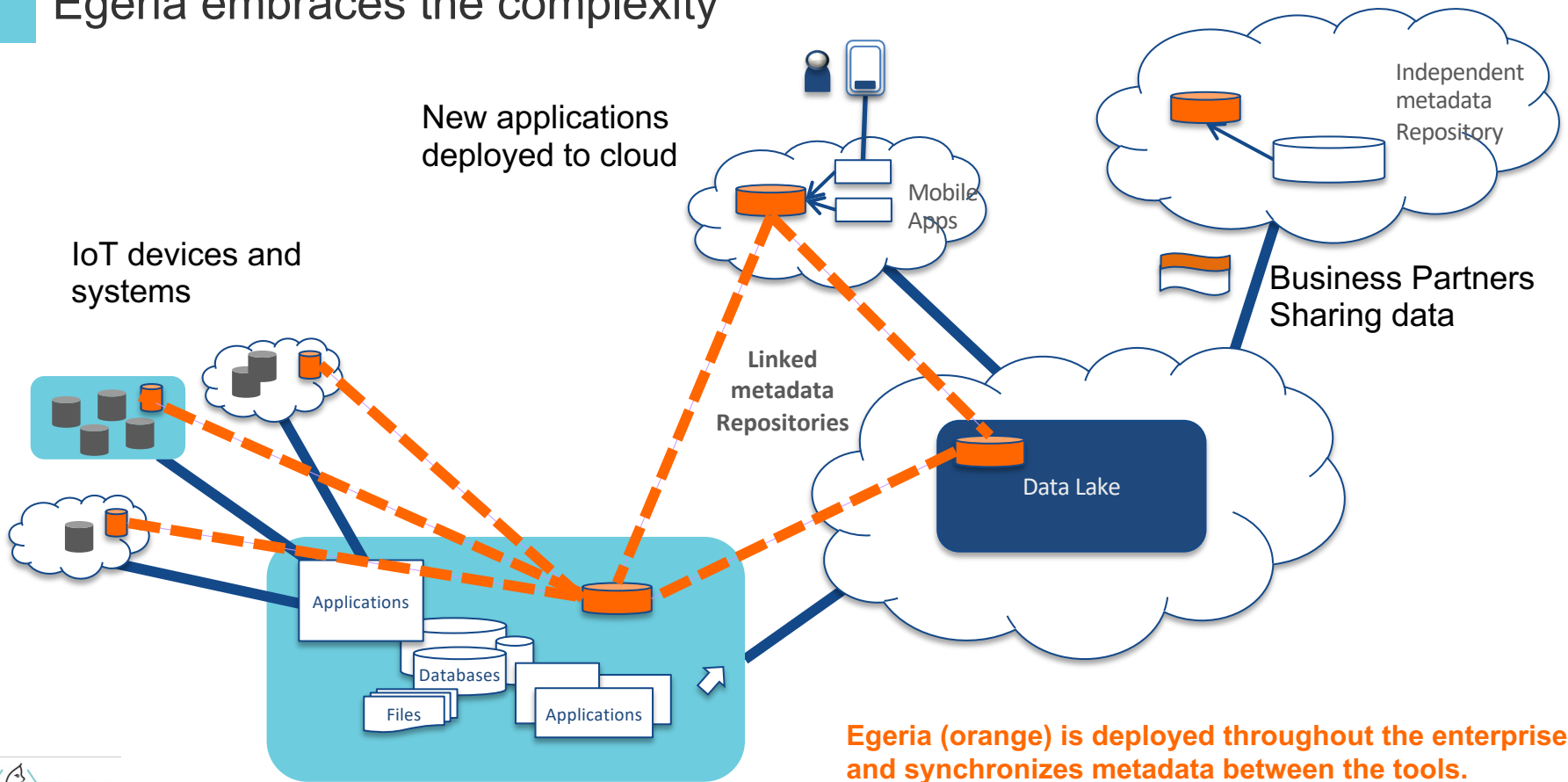
Reality check - hybrid multi-cloud world



Knowledge is separated in space and time



Egeria embraces the complexity

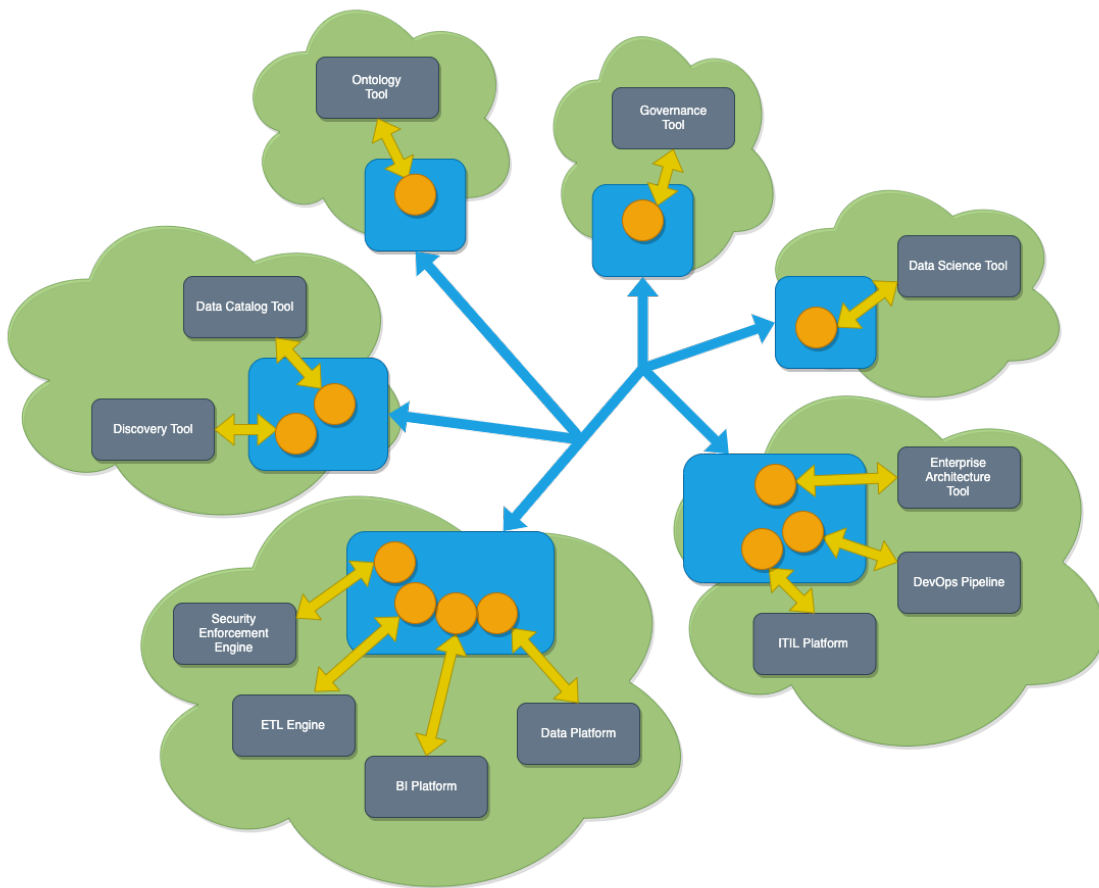


Egeria in operation

Egeria's platform (in blue) runs in each deployment environment (cloud or on premises)

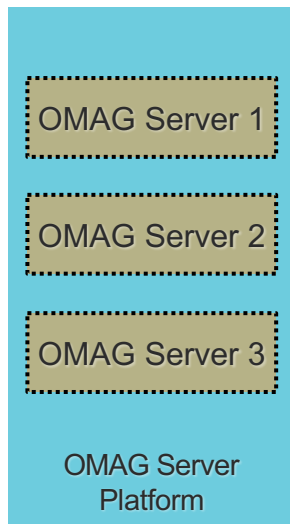
Egeria's servers (orange circles) are deployed on the platform. Each server is specialized to support the metadata needs of specific technologies.

The interaction of the servers organizes the metadata exchange

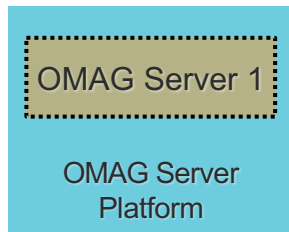


Egeria's OMAG Server Platform

↑ Open Metadata and Governance

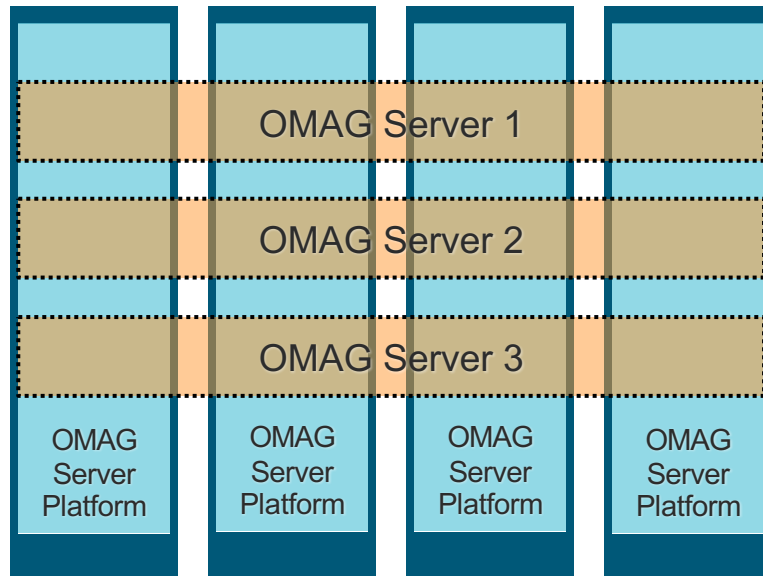


Multi-tenant

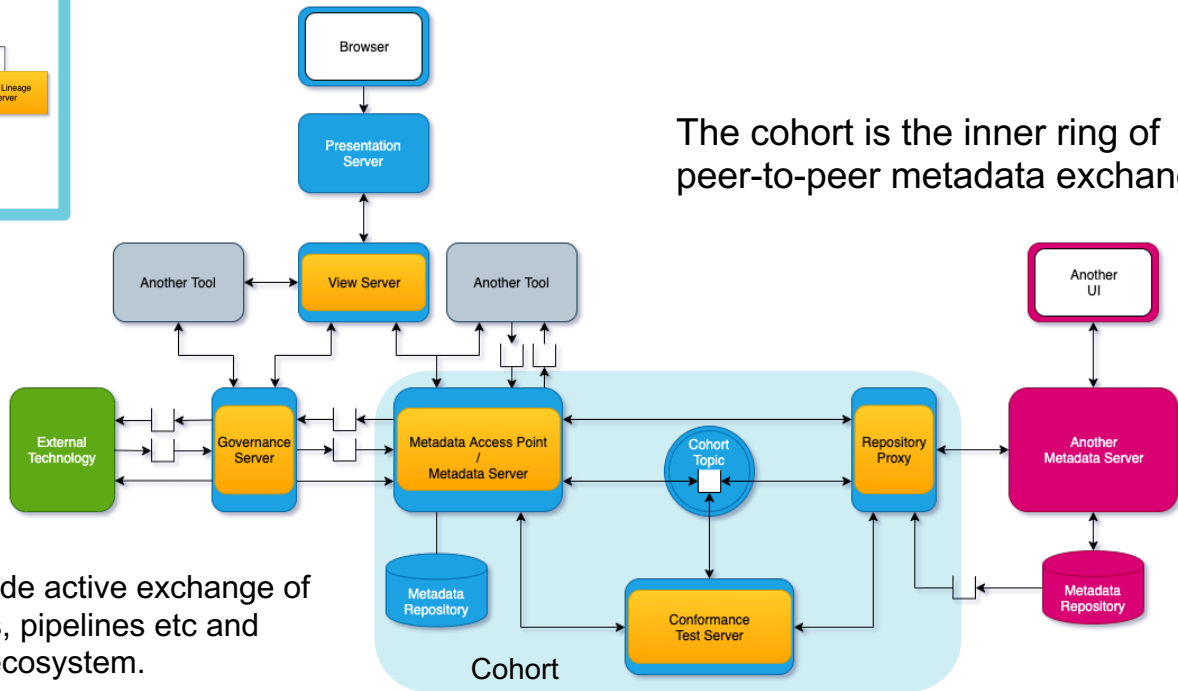
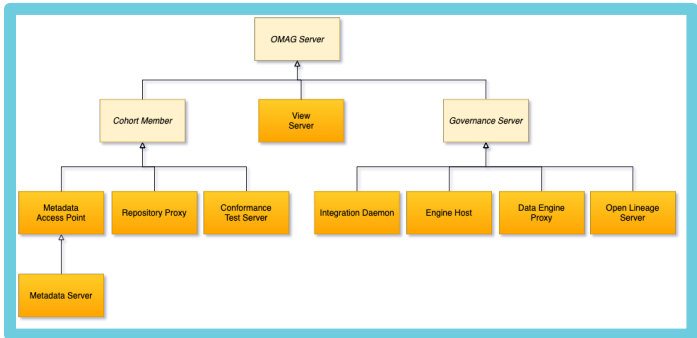


Edge

Kubernetes



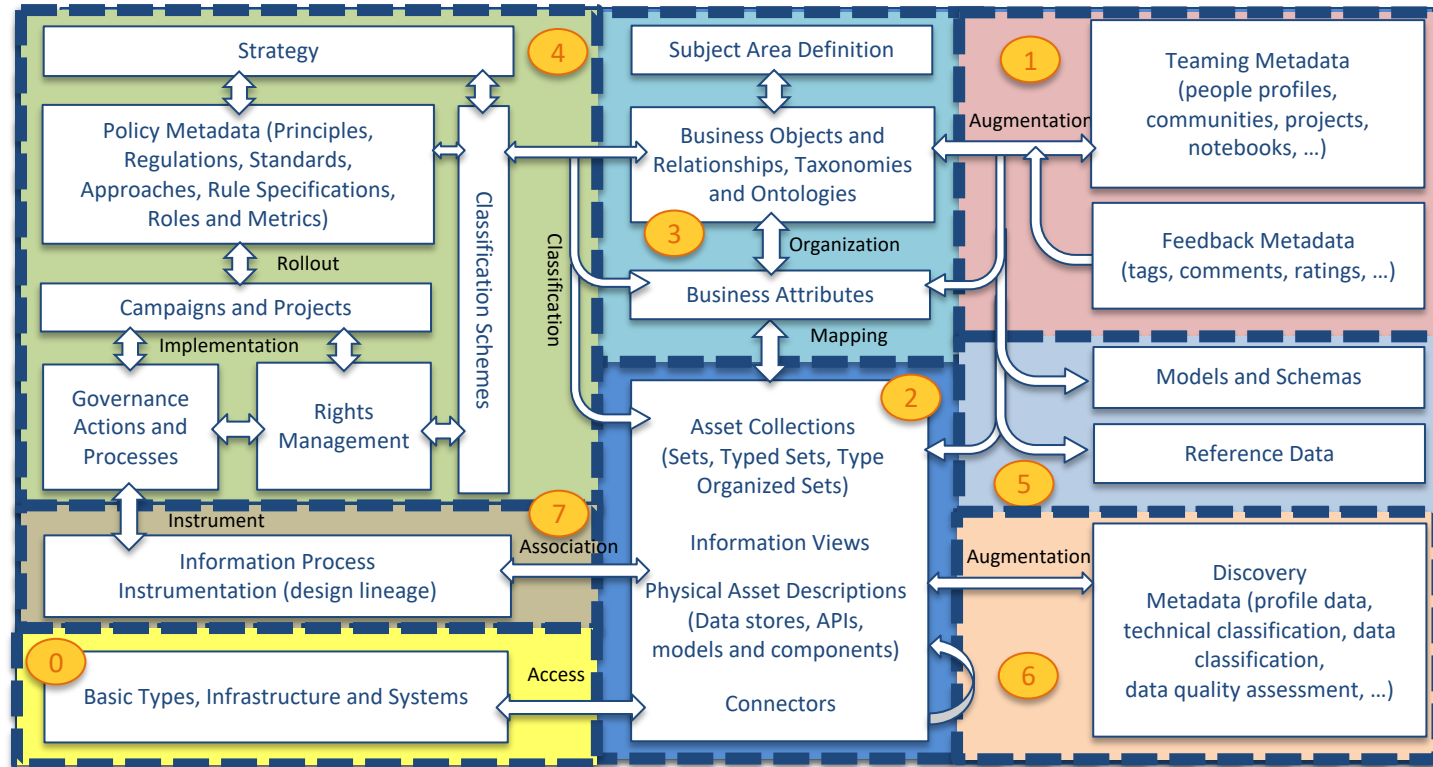
Egeria's Open Metadata and Governance (OMAG) Servers



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

Scope of metadata covered

<https://odpi.github.io/egeria-docs/types/>

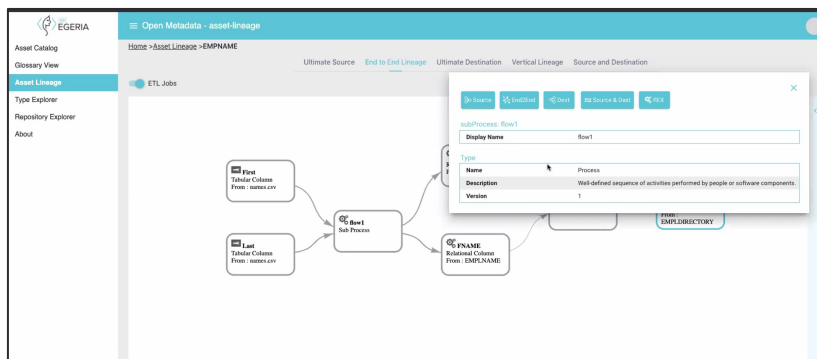


General User UI

Features

- Glossary Search
- Asset Search
- Lineage
- Repository browsing

Customizable look and feel
for Enterprises



The screenshot shows the 'Open Metadata - asset-catalog' interface. It features a sidebar with navigation options: Asset Catalog, Glossary View, Type Explorer, Repository Explorer, and About. The main area displays a table of glossary terms.

Name	Type	Description
Access Profile Name	GlossaryTerm	Access Profile Name is the Group Name
Accounting Categor...	GlossaryTerm	Accounting Category Name is the word
Accounting Structur...	GlossaryTerm	Accounting Structure Item Name is the
Agreement Alternat...	GlossaryTerm	Agreement Alternative Name is label
Agreement Alternat...	GlossaryTerm	Agreement Alternative Name Value
Agreement Descript...	GlossaryTerm	Agreement Descriptor Name identifies an ...
Agreement Display ...	GlossaryTerm	Agreement Display Name is name of the A...
Agreement Name	GlossaryTerm	Agreement Name is a label given to a speci...
Agreement Nick Na...	GlossaryTerm	Agreement Nick Name is the alternate na...
Birth Last Name	GlossaryTerm	Birth Last Name identifies the Last Name L...
Birth Last Name Prefix	GlossaryTerm	Birth Last Name Prefix is the part of a nom...

The screenshot shows the 'Open Metadata - asset-catalog' interface with a 'Details' view for a glossary term. It features a sidebar with navigation options: Asset Catalog, Glossary View, Type Explorer, Repository Explorer, and About. The main area displays a table of properties and attributes for the selected term.

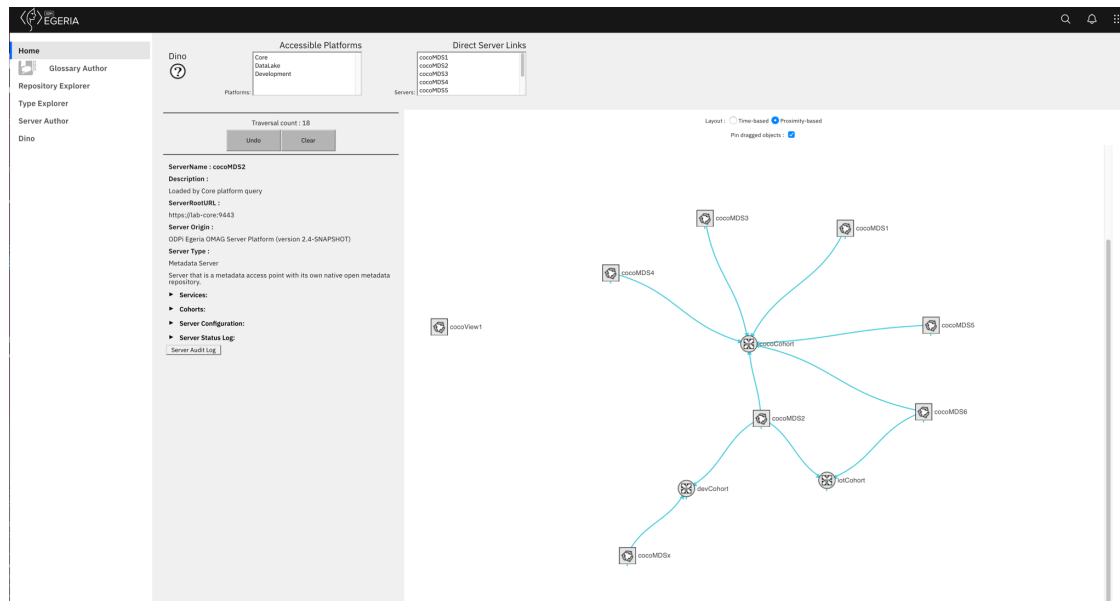
Properties	
Summary	Agreement Name is a label given to a specific Agreement. The Financial Institution or their customers often require a name for an Agreement to more easily distinguish it from other Agreement.
Display Name	Agreement Name
Qualified Name	(category)=ING Esperanto:(category)=ING Esperanto Terms:(category)=Agreement Data:(term)=Agreement Name

Type	
Name	GlossaryTerm
Description	A semantic description of something, such as a concept, object, asset, technology, role or group.
Version	1

Attributes	
Guid	term@98772b3f-23c3-4323-aad6b-73c93b0de1a4f6662c0f2e1b1ec5c-ka16b7krp.2chtee.vif6q.9lfj2j0661amsrca3p0
Created By	lsadmin_xj

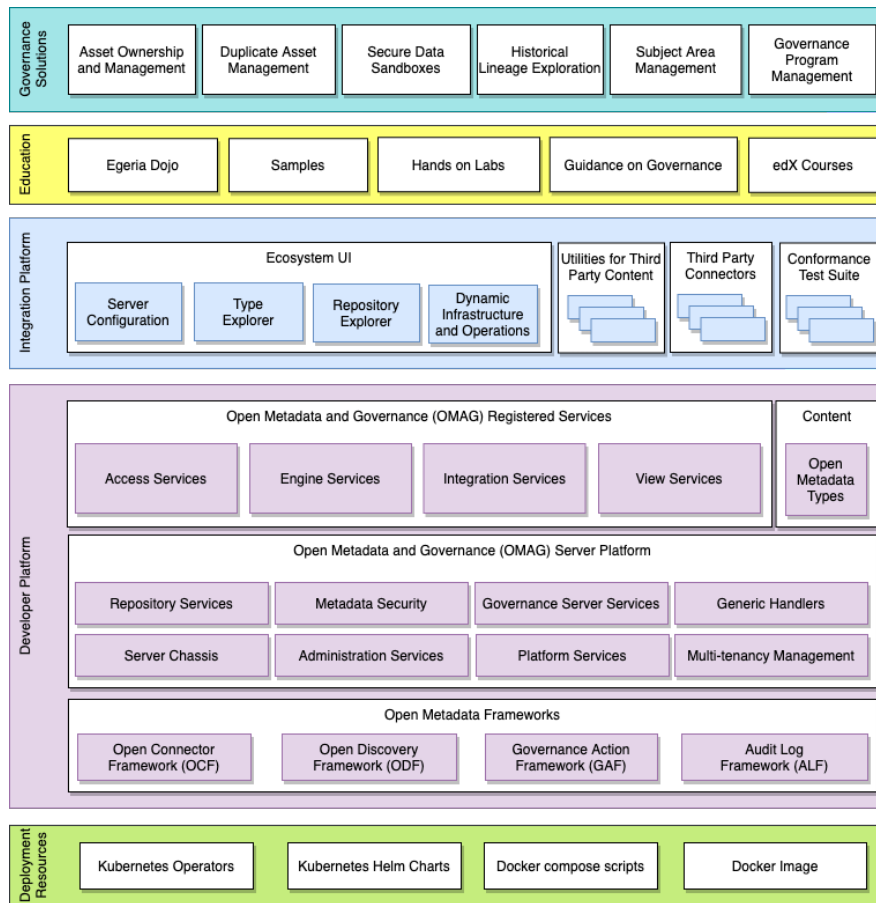
Open Metadata Ecosystem UI

- Multi-tenant UI
- Designed to support governance and open metadata teams



Egeria's full stack

- Egeria is more than code
 - Deployment
 - Education
 - Content packs
 - Solution cookbooks
 - Conformance test suite



How does the Egeria community work

- Iteratively perform a series of practical experiments that ignore the boundaries of traditional approaches.
 - Holistic approach
 - Open, public collaboration
 - Developing interest and partnership as we go
 - Agile development with feedback loop
 - Enterprise grade software

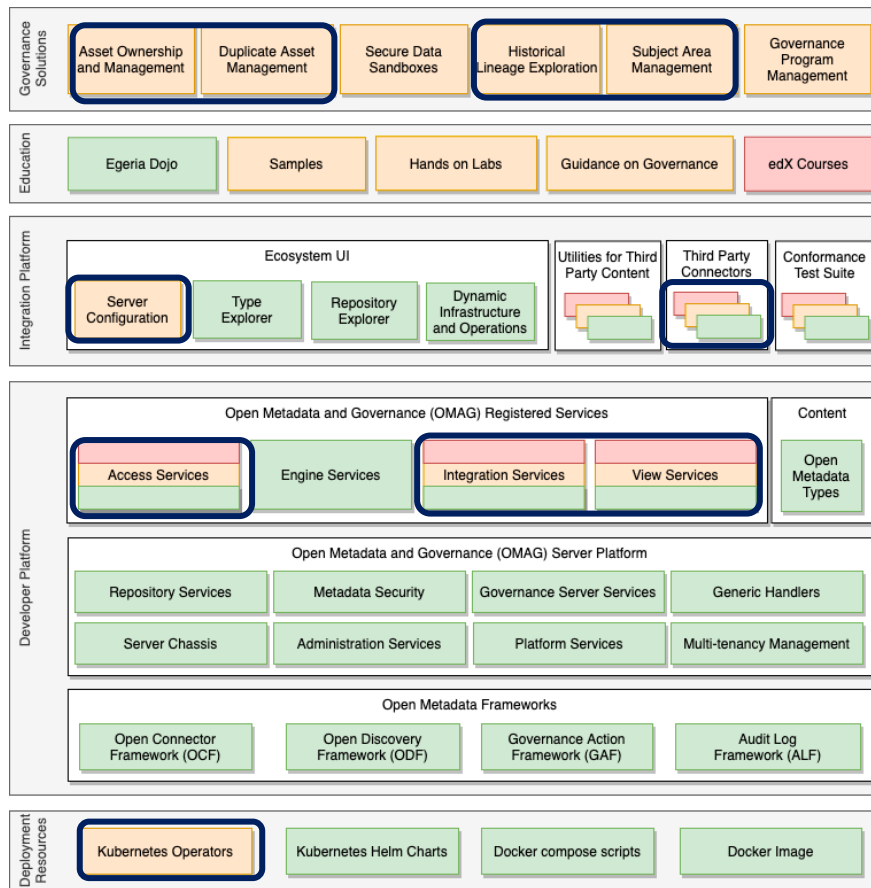
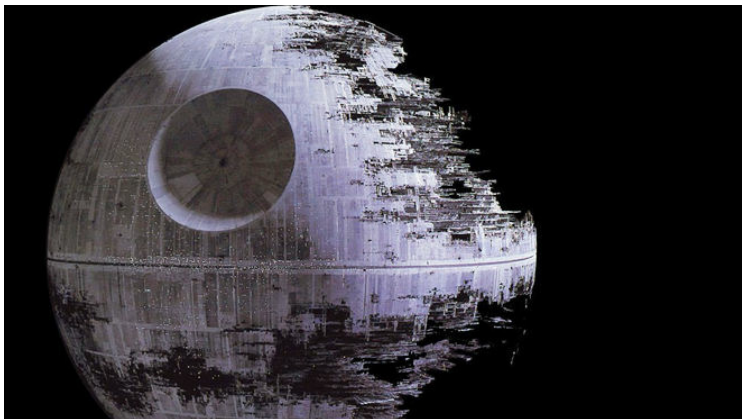


Why do vendors work with Egeria?

- Linking old platform with new platform
- Linking multiple versions of own product deployed across an enterprise
- Offload integration implementation costs
- Access to third party metadata:
 - open source technology's and other vendor's metadata
- Governing third party technologies

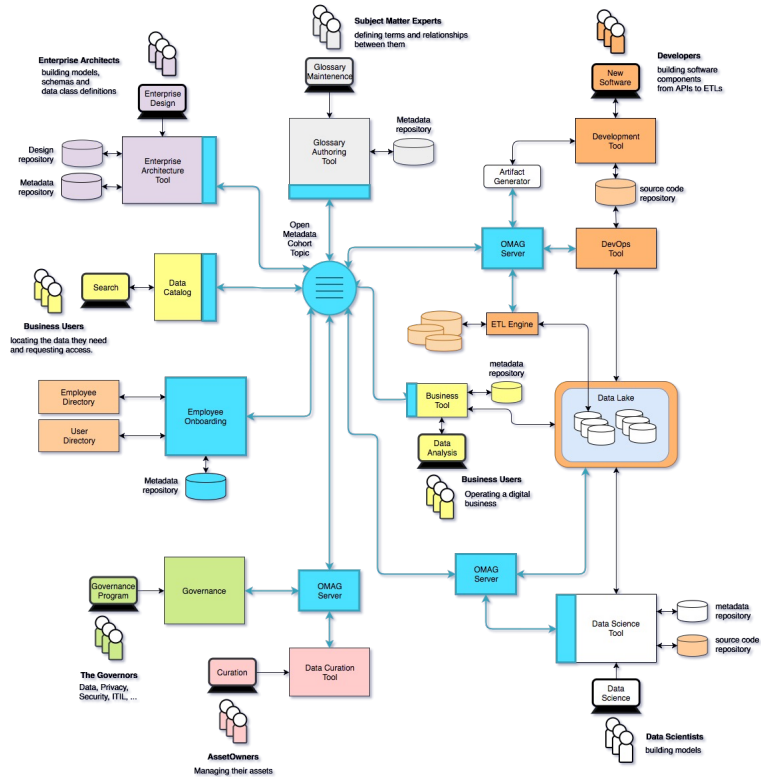
Development Status

- Development in plain sight
- Monthly releases
 - What is ready, goes ...



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



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 heterogeneous 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