

EKGF

The Path from Knowledge Graph to Enterprise Knowledge Graph

presented May 3, 2021

Agenda

- Welcome – Jeannine Wisnosky Stehlin
- From Tables to Ontologies & Knowledge Graphs – Dennis Wisnosky
- The Enterprise Knowledge Graph Foundation – Dennis Wisnosky
- Meet EKG Foundation Founding Members – Jeannine Wisnosky Stehlin
- Differences between KG and EKG – Jacobus Geluk
- The EKG Principles, the nucleus of the EKGF Manifesto – Jacobus Geluk
- EKG Maturity Model Status – Jacobus Geluk & Pete Rivett
- Quick intro to the EKGF Collaboration Process – Pete Rivett
- Requirements for EKGF Portals: interactive session – Pete Rivett
- Q & A
- Meet us in the lounge

About EKG Foundation

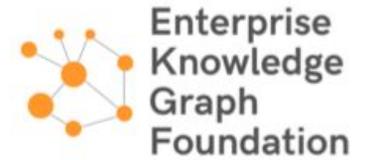
The Enterprise Knowledge Graph Foundation (EKGF) was established in 2020 as a non-profit trade association to define best practices and mature the marketplace for Enterprise Knowledge Graph (EKG) adoption, including:

- **STANDARDS**
Adopt semantic standards and define additional standards for large-scale secure, resilient & interoperable EKGs
- **METHOD**
Develop best practices for accelerated EKG deployment
- **ARTIFACTS**
Curated repository of reusable use cases, models, ontologies, datasets and other resources
- **COMMUNITY**
Build the mechanism for engagement and shared knowledge
- **ADVOCACY**
Advance the business cases for EKG adoption

EKG Foundation Organizing Team

- **Dennis Wisnosky**
Founder of the Wizdom Companies and former CTO and CA (business), US Department of Defense
- **Jacobus Geluk**
CEO, agnos.ai
- **Pete Rivett**
CKO, agnos.ai and Architecture Board member of the Object Management Group
- **Michael Atkin**
Founder and former CEO of the EDM Council
- **Jeannine Wisnosky Stehlin**
Nonprofit Managing Director

EKG FOUNDATION



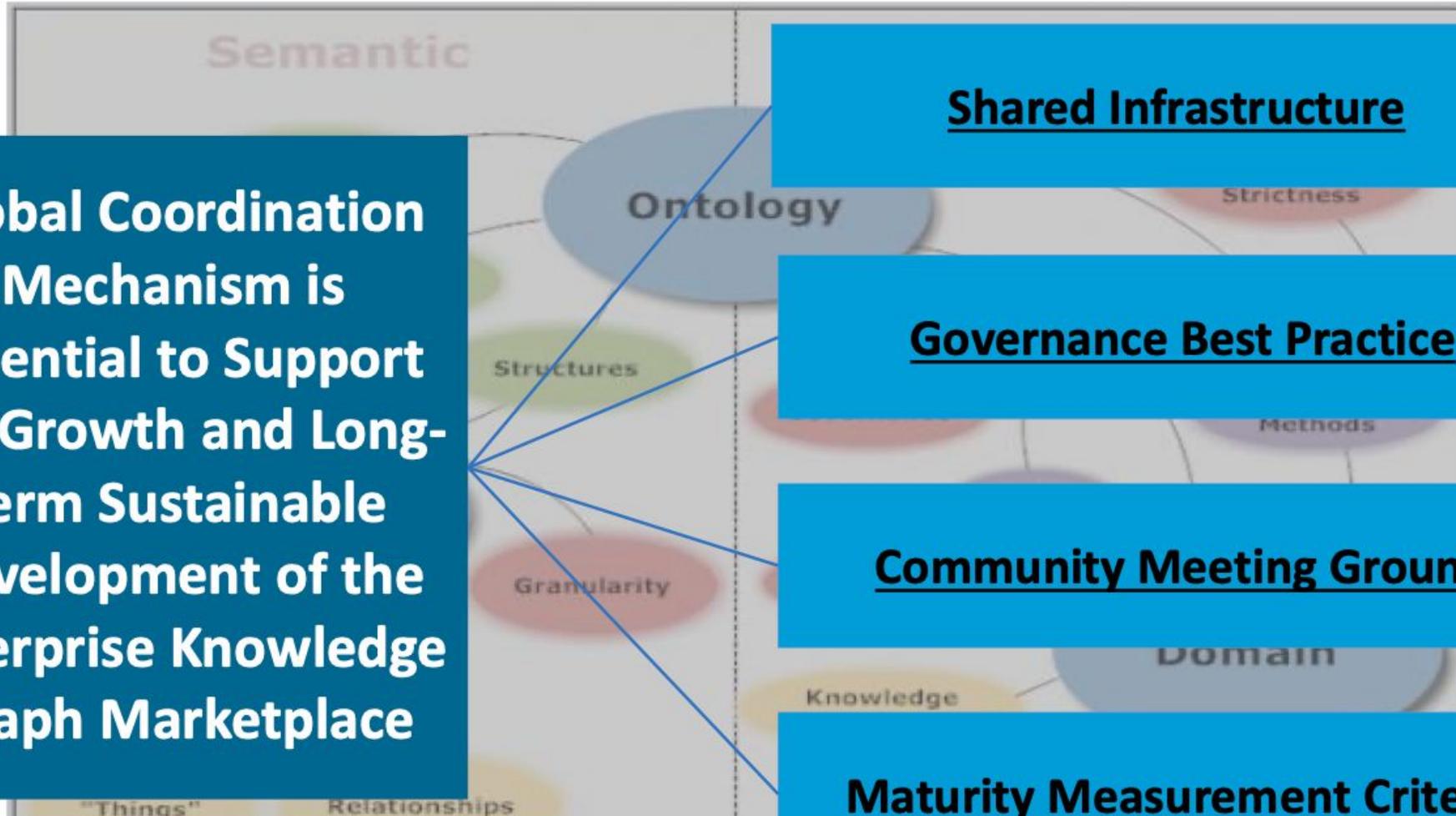
Global Coordination Mechanism is Essential to Support the Growth and Long-Term Sustainable Development of the Enterprise Knowledge Graph Marketplace

Shared Infrastructure

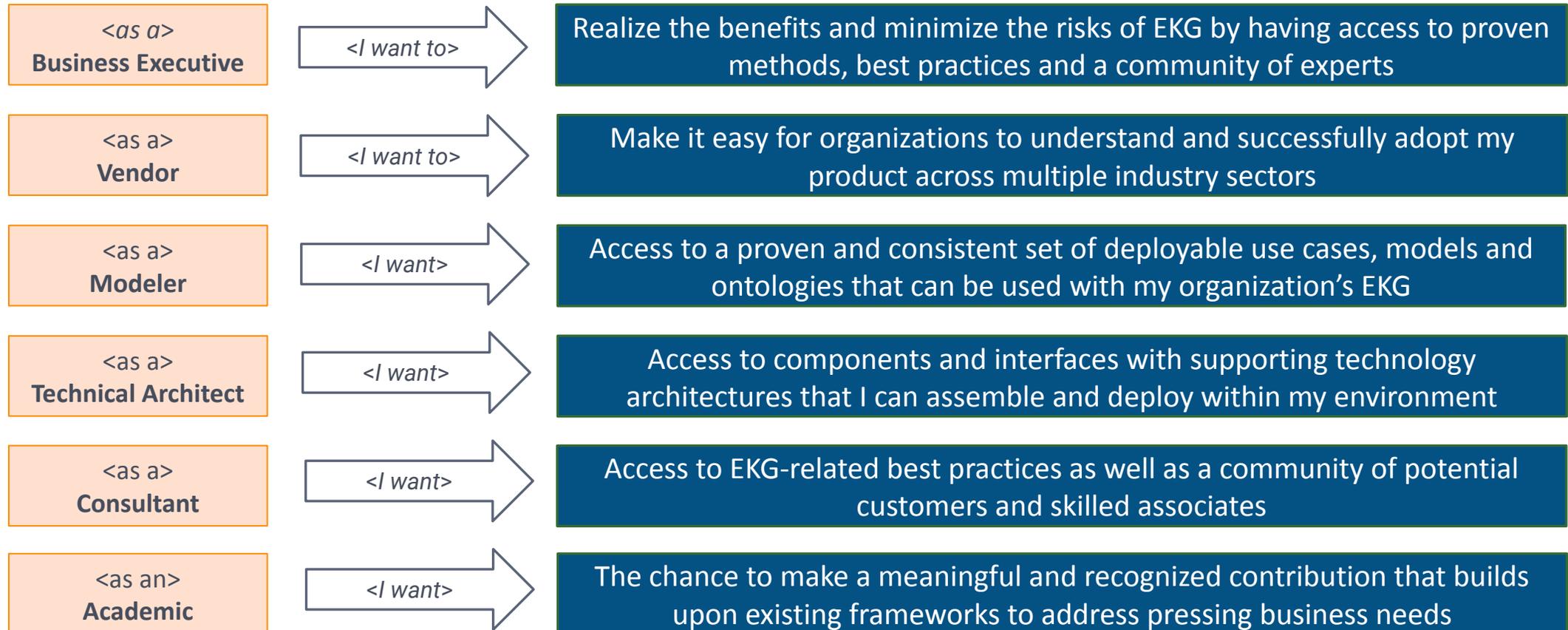
Governance Best Practice

Community Meeting Ground

Maturity Measurement Criteria



Stakeholder User Stories



Membership Categories

End Users*

semantic technology for internal use

Service Providers*

vendors, consultants and publishers

Affiliates

academia, government, press and non-profits

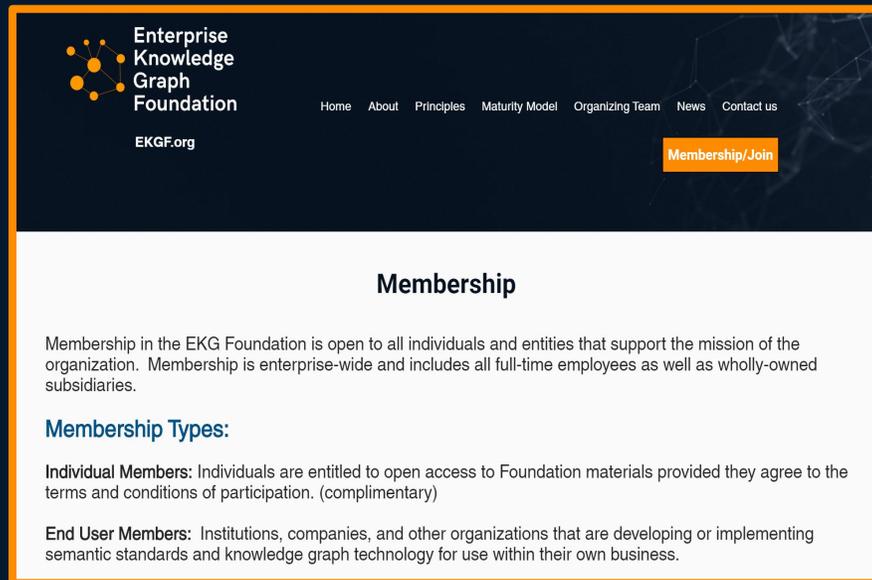
Individuals

complimentary access

* Voting rights

Next Steps From You...

1. Join the Foundation



The screenshot shows the EKG Foundation website's membership page. The header includes the logo, navigation links (Home, About, Principles, Maturity Model, Organizing Team, News, Contact us), and the URL EKG.org. A 'Membership/Join' button is visible in the top right. The main content area is titled 'Membership' and contains a paragraph explaining that membership is open to all individuals and entities supporting the mission. Below this, there is a section for 'Membership Types' with two sub-sections: 'Individual Members' and 'End User Members', each with a brief description of their benefits and requirements.

Enterprise Knowledge Graph Foundation
EKG.org

Home About Principles Maturity Model Organizing Team News Contact us

Membership/Join

Membership

Membership in the EKG Foundation is open to all individuals and entities that support the mission of the organization. Membership is enterprise-wide and includes all full-time employees as well as wholly-owned subsidiaries.

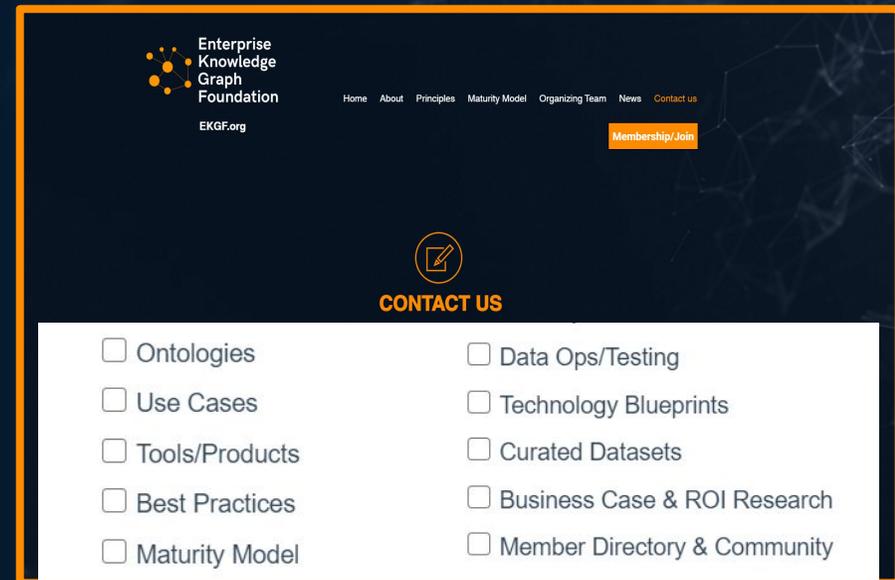
Membership Types:

Individual Members: Individuals are entitled to open access to Foundation materials provided they agree to the terms and conditions of participation. (complimentary)

End User Members: Institutions, companies, and other organizations that are developing or implementing semantic standards and knowledge graph technology for use within their own business.

<https://www.ekgf.org/membership>

2. Register Your Priorities



The screenshot shows the EKG Foundation website's contact page. The header is identical to the membership page. The main content area is titled 'CONTACT US' and features a list of priority areas, each with an unchecked checkbox. The areas are arranged in two columns: Ontologies, Use Cases, Tools/Products, Best Practices, and Maturity Model on the left; and Data Ops/Testing, Technology Blueprints, Curated Datasets, Business Case & ROI Research, and Member Directory & Community on the right.

Enterprise Knowledge Graph Foundation
EKG.org

Home About Principles Maturity Model Organizing Team News Contact us

Membership/Join

CONTACT US

- Ontologies
- Use Cases
- Tools/Products
- Best Practices
- Maturity Model
- Data Ops/Testing
- Technology Blueprints
- Curated Datasets
- Business Case & ROI Research
- Member Directory & Community

<https://www.ekgf.org/contact>

EKGF Founding Members



Global IDs provides software for enterprise information management (EIM).



Stardog, the leading Enterprise Knowledge Graph platform, turns data into knowledge to power more effective digital transformations.



eccenca is a leading provider of enterprise knowledge graph management software and solutions.



data.world makes it easy for everyone—not just the "data people"—to get clear, accurate, fast answers to any business question.



Ontotext is a global leader in enterprise knowledge graph technology and semantic database engines.

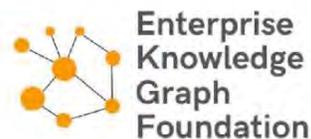


agnos.ai is an Enterprise Knowledge Graph consultancy assisting its clients in their EKG journey from strategy to production.



THE SMART DATA COMPANY®
Cambridge Semantics provides Anzo, the scalable knowledge graph platform for data integration and analytics.

Founding Members



Wizdom is a recognized leader in providing our clients with innovative process-based business solutions for improving performance.

EKGF Member – agnos.ai



[agnos.ai](#) is a specialist consultancy that assist its clients in their EKG journey, transforming their business.

We are an international partnership of EKG engineers & architects, methodology specialists, software engineers, and data management experts.

We understand the challenges of managing data in complex and interconnected environments.

We help our clients to deliver their strategic EKG use cases all the way to production.

EKGF Member — Cambridge Semantics



[Cambridge Semantics Inc.](#) is a modern data management and enterprise analytics software company. Our product Anzo[®] is a scalable knowledge graph platform for modern data integration and analytics. Anzo dramatically simplifies and accelerates the integration, modeling, and blending of siloed data into insight-rich knowledge graphs at enterprise scale. Anzo is built on AnzoGraph[®], the fastest & most scalable knowledge graph engine supporting data integration, graph algorithms, data warehouse-style analytics, feature engineering for Machine Learning, and more. The company supports solutions that enable IT departments and business users across Life Sciences, Financial Services, Government, Manufacturing, and other industries to accelerate data delivery and provide meaningful insights across the organization. Cambridge Semantics is based in Boston, Massachusetts.

About Cambridge Semantics

Scalable knowledge graphs for modern data integration and analytics.

- Based in Boston
- Origins in IBM and Netezza
- Anzo 5.0 GA Feb 2020
- Featuring enterprise-scale OLAP graph database engine

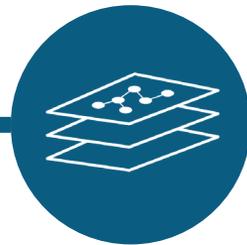


A **scalable knowledge graph platform** for modern data integration and analytics

Anzo connects and models related data in a real-world representation of data at scale, surfacing new insights and fueling pervasive analytics.



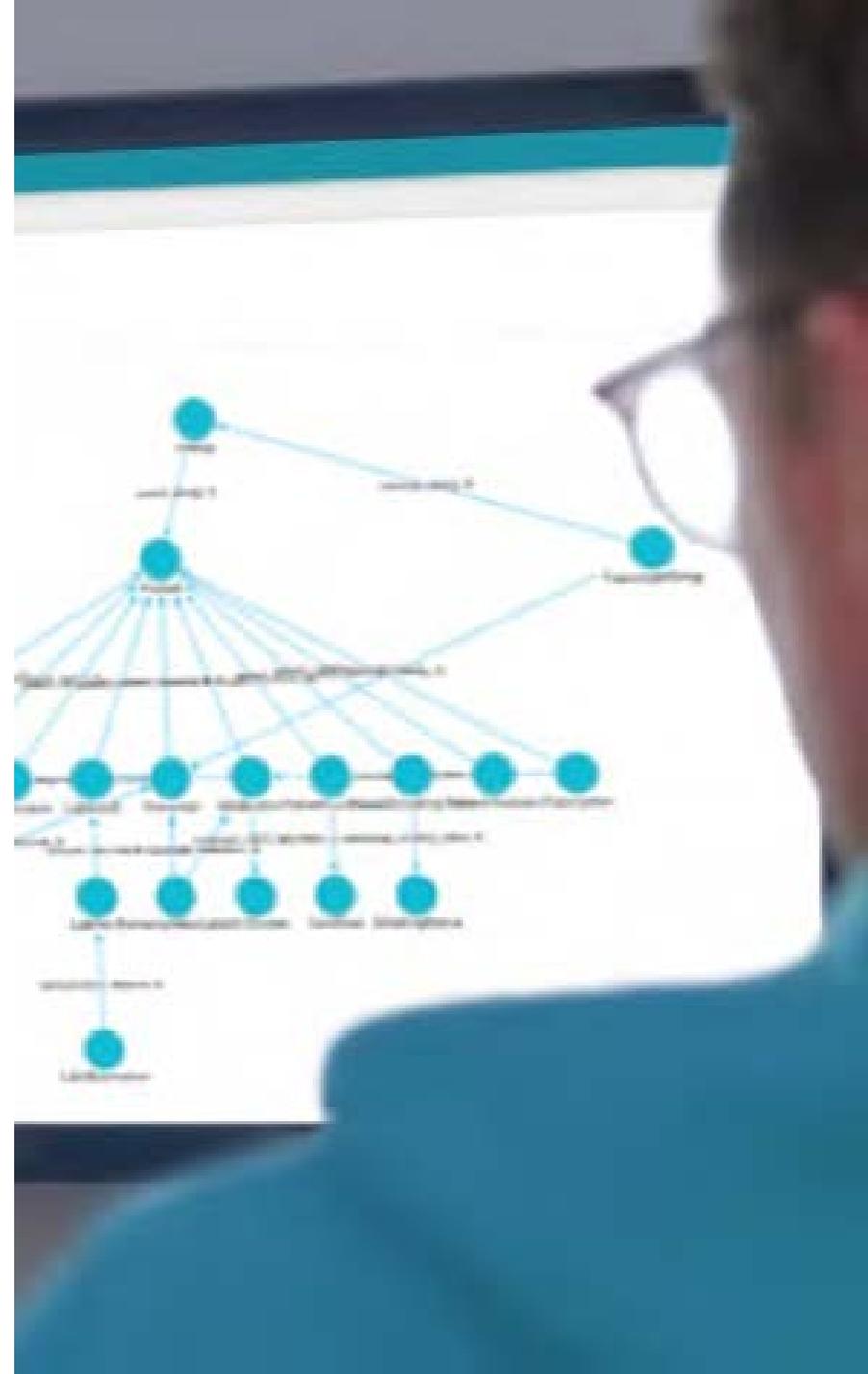
Knowledge Graph Management and Metadata Cataloging



AnzoGraph MPP OLAP Knowledge Graph Engine



Enterprise-grade cloud deployment and security



EKGF Member — data.world



data.world's Knowledge Graph as a Service connects your data and knowledge together so you can build data products faster. Our integrated platform gives you tools to design the domain model, declaratively map it to your physical sources, and construct the resulting graph. Organizations first use the data.world data catalog to capture and organize their data ecosystem and then model and apply their understanding of the business against it with Knowledge Graph as a Service. Bring our semantic layer, services, and expertise to your data ecosystem to build and scale data products like never before.

EKGF Member — eccenca



eccenca is a leading provider of enterprise knowledge graph management software and solutions. Its flagship product Corporate Memory helps companies to master the complexity and dynamics of their knowledge, product and device lifecycles. The eccenca platform enables companies to manage rules, constraints, configurations and expert knowledge as well as data in one central application. This empowers customers to digitally scale knowledge for automation.

EKGF Member — Global IDs



Global IDs provides software for enterprise information management (EIM). Over the last 14 years, Global IDs' products have been implemented in some of the world's largest companies. Our core product, the Global IDs Data Ecosystem Management Suite addresses the problem of managing complex data. In contrast to traditional data management software that focus on data silos, Global IDs focuses on data ecosystem problems that rely on a holistic understanding of enterprise data.

EKGF Member — Ontotext



Ontotext is a global leader in enterprise knowledge graph technology and semantic database engines. We connect the dots of your enterprise knowledge. By leveraging AI technologies, we help enterprises get a competitive advantage in dynamic environments. We fuse and smarten up their proprietary information using global knowledge as context for interpretation and source for enrichment.

Products and services: Our most popular product is GraphDB - semantic graph database engine. We also offer end-to-end solution We do that by crafting big knowledge graphs to enable unified data access and cognitive analytics. We link diverse data, enrich it via text analysis and index it in GraphDB for semantic search.

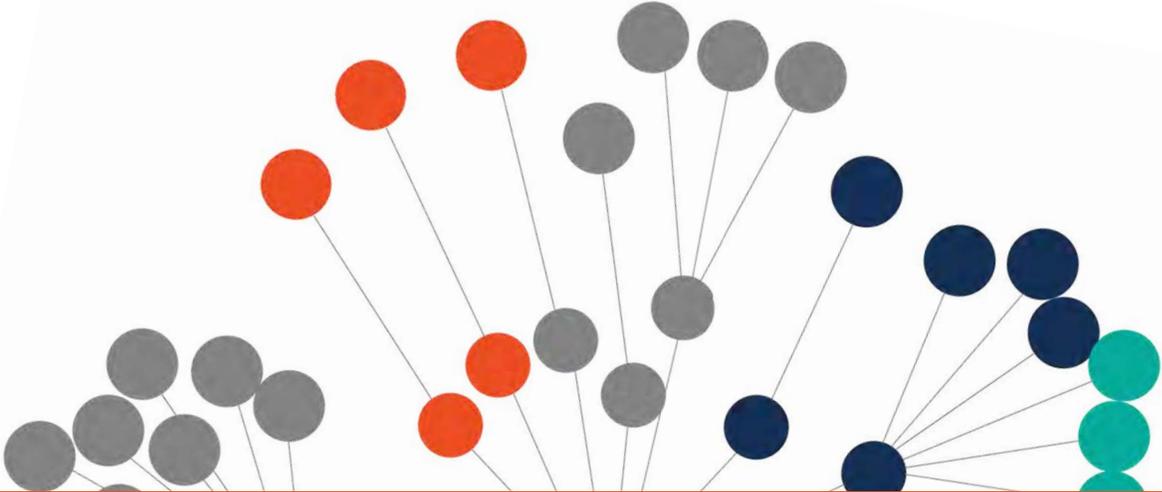
Some our customers: S&P, BBC, FT, Elsevier, Top-5 US Bank, UK Parliament, NASA, Fujitsu, Novartis, PSA

Ontotext joined EKG Foundation because: While KG are in high demand among the thought leaders in different industries, there is much more work to be done to set the expectations among business sponsors, enterprise architects and even technology analysts. The best practices, the resources and the methodologies that EKG will develop and gather is very important in this process. Ontotext will contribute to the EKG's activities its extensive experience in moving this technology from early adoption to mass market acceptance



ontotext

making sense of text and data



Knowledge Graphs as Hub for Data, Metadata and Content

We help enterprises get profound insights via linking, analysis and exploration of:

- Diverse databases
- Text documents and other content
- Proprietary & Global data

April 2021

Ontotext Introduction

○ Leader

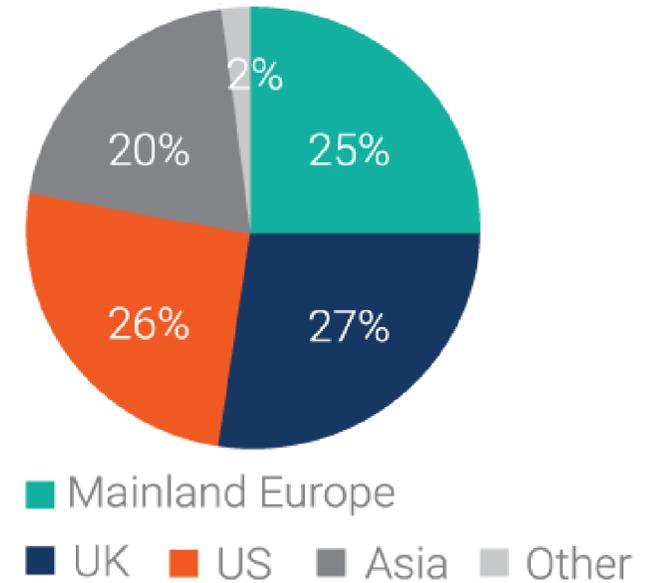
- ✓ Semantic technology vendor established year 2000
- ✓ Knowledge graphs thought leader

○ Profitable and growing

- ✓ **Global:** 80% of revenue from multinational companies
- ✓ **Clients:** S&P, BBC, FT, Elsevier, Top-5 US Bank, UK Parliament, NASA, Fujitsu, Novartis, PSA

○ Innovator

- ✓ Member of W3C, ODI, STI, EKGf, Linked Data Benchmarking Council, DBpedia Association, ...
- ✓ 30+ collaborative research projects with the best academic groups in Europe



EKGF Member — Stardog



Stardog, the leading Enterprise Knowledge Graph platform, turns data into knowledge to power more effective digital transformations. Industry leaders including BNY Mellon, Bosch, and NASA use Stardog to create a flexible data layer that can support countless applications. With Stardog, customers reduce data preparation timelines by up to 90%. Stardog is a privately held, venture-backed company headquartered in Arlington, VA. For more information, please visit www.stardog.com.



STAR·DOG

Stardog queries all on-prem and cloud data silos directly so you don't have to move or copy data in order to make data-driven decisions.

We help companies:

- Build innovative products
- Create efficient, data-driven operations
- Unlock access to data

Achieving the best results through:

- Virtualization
- Semantic Graph
- Inference Engine



SPRINGER NATURE



SIEMENS

ELSEVIER

NOKIA



EKGF Member — Wizdom



Wizdom is a recognized leader in providing our clients with innovative process-based business solutions for improving performance. Wizdom's integrated training and consulting gives you an improved enterprise-wide business process and enables you to communicate, plan, and execute a unified corporate strategy. We help companies grow by aggressively pursuing new opportunities and employing the best business practices.

The EKG Principles

1. Identity

Unique, opaque, lifetime resolvable EKG/IRIs

2. Meaning

Machine-readable meaning

3. Distributed

EKG objects can exist anywhere

4. Open World

Fundamentally support any version of the truth

5. Self-describing

EKG-deployable datasets, enforceable policies

6. Measurement

Comprehensive built-in measurements

7. Business Orientation

Modular structure of reusable business use cases

8. Control

Executable models for entitlements, privacy and business policies

9. Ecosystem

Interoperable graphs anywhere

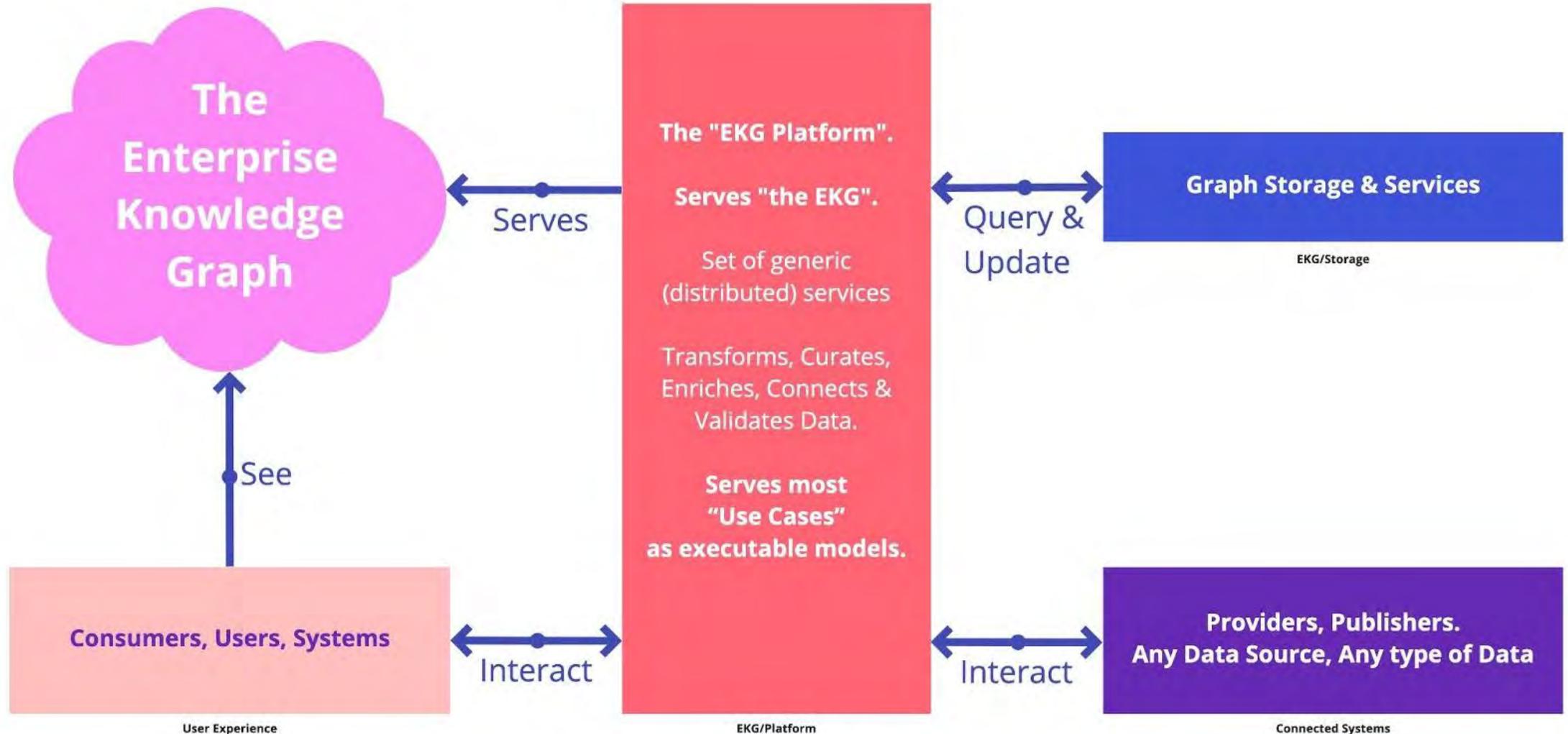
10. Standards

Standards for data, ontologies and EKG interfaces

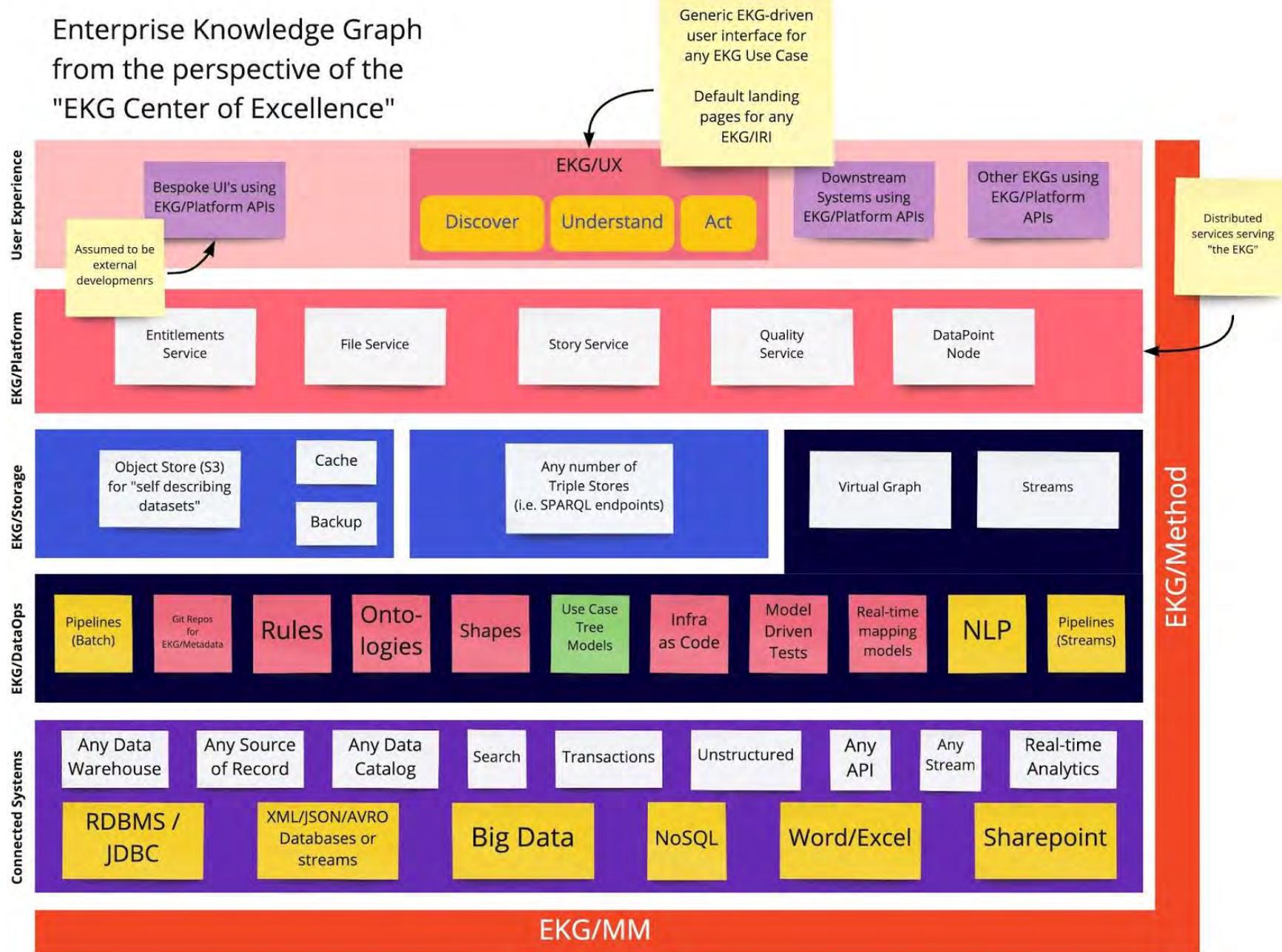
Paradigm Changes

- Open World
- Graph, Semantic Graph, Smart Data
- Linked Data (for the enterprise)
“All data connected”
- DevOps, DataOps, TDD/BDD
- IaC, GitOps
- Continuous Deployment
- Agile
- Use Case Trees, “Strategic Use Cases”
- Cloud, Containers, Micro-services, IoT, Edge Computing
- Digital Transformation
- Data fabric / Knowledge fabric,
“no (more) silos”
- Digital Twins
- “Machine readable meaning”,
Ontologies, Shapes
- ML/AI, AI-as-a-service, EKG driven AI
(smart AI?)

EKG vs EKG Platform



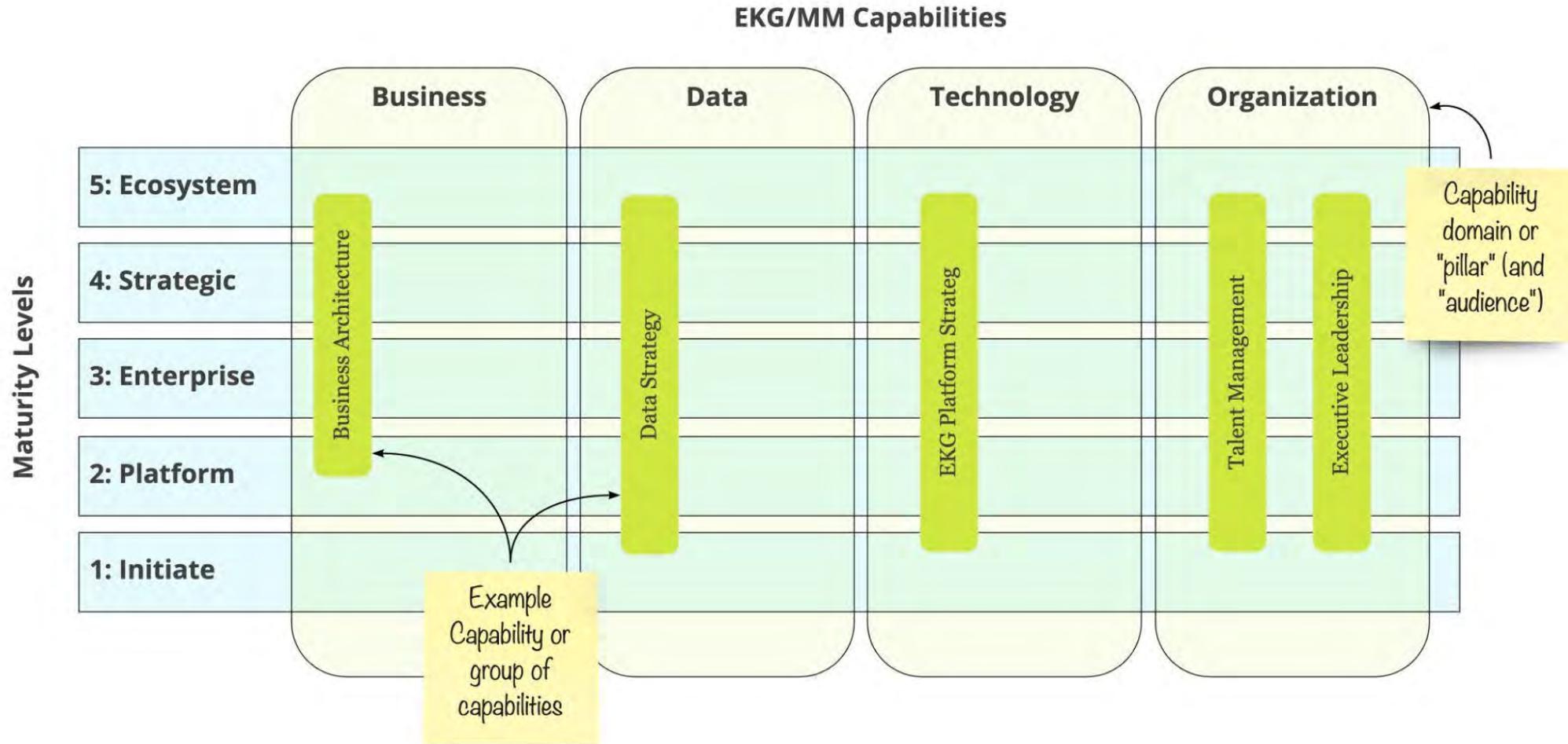
Enterprise Knowledge Graph from the perspective of the "EKG Center of Excellence"



EKG/MM Status

- The foundation decided to make it all open, public git repo.
- Community around the maturity model is growing.
- Participants are contributing.
- “Version 1.0” will cover all major capabilities and will be published this year.
- Maturity model has many purposes but primary aim is to have a structured way to explain the EKG vision to the four key audiences and manage expectations.
- Version 2.0 will add detailed scoring criteria.
- Version 3.0 will be fully EKG-based, including ontologies for each capability area.

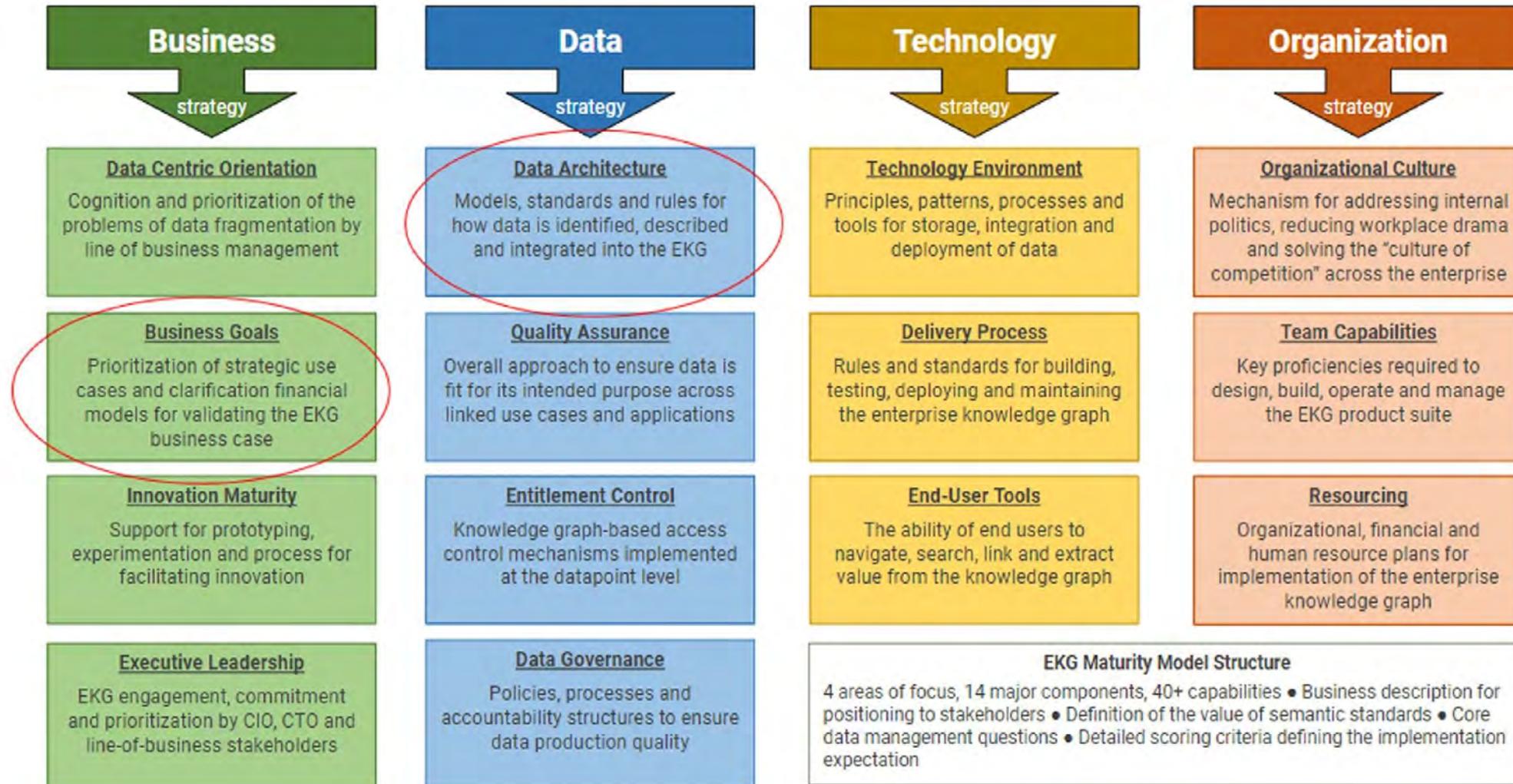
EKG/MM Structure



EKG/MM Current focus areas

Pillar	Component	Capability	Measurable ability
A Business	A.1 Business Strategy Actuation	A.1.1 Business Vision	Linking EKG to strategic business and organizational objectives
		A.1.2 Business Goals	Prioritization of strategic use cases and financial models
		A.1.3 Business Tactics	Incremental steps to address the challenges of data fragmentation
	A.2 Business Architecture Elaboration	A.1.4 Operating Model	Performance-improving journeys that cut through organizational silos
		A.2.1 Market Segmentation	Shared understanding of markets, products, and clients
		A.2.2 Value Chain	Alignment of business capabilities with EKG pathway
	A.3 Business Enablers	A.2.3 Change Management	Adopt alternative organizational and operational approaches
		A.3.1 Performance Management	Measure a process to manage performance
		A.3.2 Risk Management	Identify & assess potential risks and create mitigation plans
B Data	B.1 Data Strategy	A.3.3 Supply Chain Management	Managed flow of goods, services & information between businesses
		B.1.1 Goals & Objectives	EKG as the semantic data fabric for the organization
		B.1.2 Business Case	Business rationale, justification and Return On Investment (ROI) of the EKG
	B.2 Data Architecture	B.2.1 Identity Resolution	Standards for ensuring content is identified and resolvable
		B.2.2 Ontologies & Data Models	Standards and tools for ensuring the shared meaning of data
		B.2.3 Inventory Management	Cataloguing content, sources, provenance, and lineage in the EKG
	B.3 Data Quality	B.2.4 Business Terminology	Process for ensuring that glossaries are linked to ontologies
		B.2.5 Data Integration	Patterns and approaches for moving and transforming data
		B.3.1 Framework	Strategy and approach for managing data quality
		B.3.2 Business Rules	Approach and rules to validate fit-for-purpose data quality
		B.3.3 Execution	Checkpoints and control processes for managing data quality
		B.4 Data Governance	B.4.1 Operating Model
	B.4.2 Data Management Policy		Policies, procedures, and standards for managing the data lifecycle
	B.4.3 Data Production & Consumption		Management of the data production and manufacturing process
	C.1 Technology Strategy	B.4.4 Entitlement Management	Enterprise knowledge graph access control mechanisms
B.4.5 Critical Data Elements		Identification and tracking of and uses	
B.4.6 Risk & Control Environment		Creation and implementation of the data control environment	
	C.1.1 Technology Landscape	Plans for physical infrastructure, applications, and automation	
	C.1.2 EKG Technology	Platforms and tools for storage and integration	
	C.1.3 Process Rationalization	Technology considerations related to process redesign	

EKG Maturity Model Coverage



Introducing the EKG Collaboration Process

- Currently one major artifact, the EKG/MM document
- Others will get started soon
 - e.g. “EKG/Manifesto”, “EKG/Method” and “EKG/Catalog”
- Initially, just text in public git-repos, anyone can contribute
 - e.g. <https://github.com/EKGF/ekg-mm>
 - license is [CC BY-SA 4.0](#)
- However, ideally, every contributor becomes a member
 - free for individual members.
 - contributors will be named in appendices (and/or GitHub)
- Change process is based on common DevOps / agile practice
 - Documented in appendix of EKG/MM PDF
- Weekly “standup” calls per workgroup

How we will work

Collaboration: mutual contribution, information sharing, open source code, consensus priorities (driven by members)

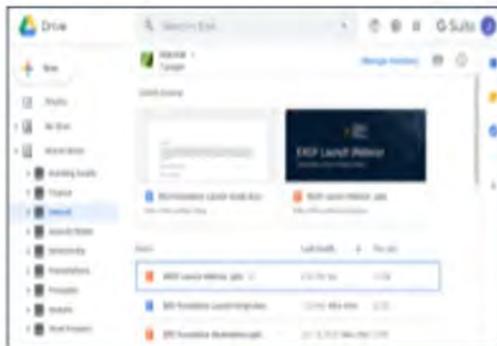
Github - open source code



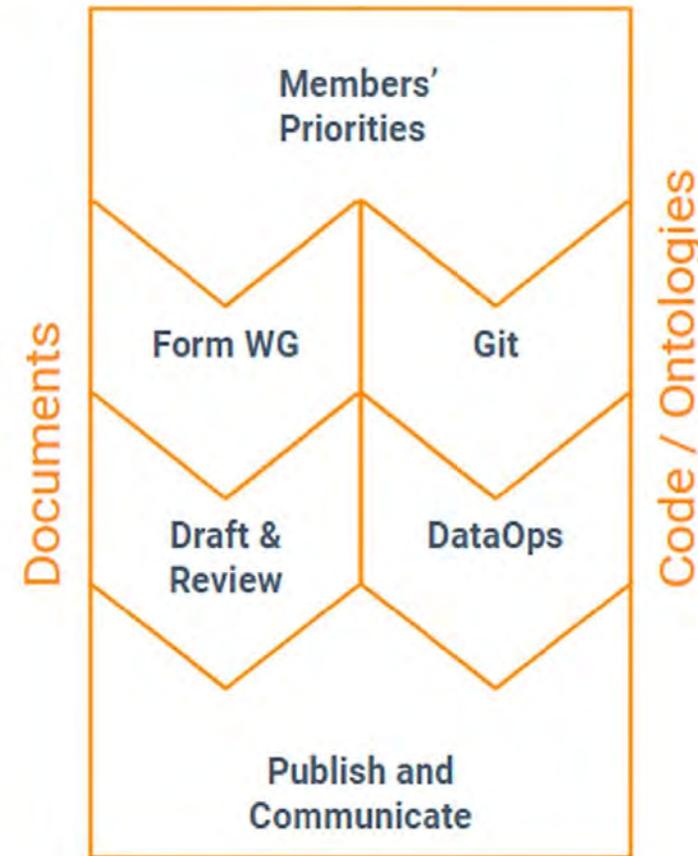
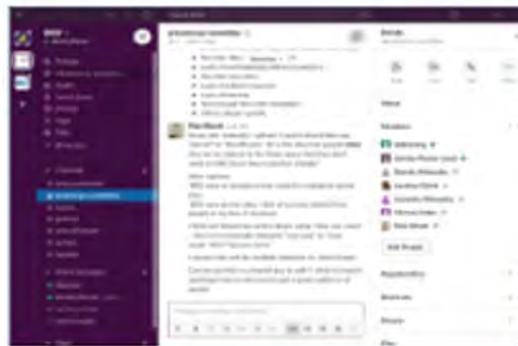
Web - publishing & portals



Google Drive - documents



Slack - member communication



EKGF Portals

Initial Portals & Audiences

Use Cases

Executive, Vendor

- Understand the business problems that can, and have, been addressed;
- Use a framework for structuring projects
- Build your own use case tree from EKG content
- Get directed to relevant reuse points in other portals (e.g. for ontologies)

Best Practices

*Project Lead,
Consultant*

- Reduce the risk and cost associated with new EKG use cases
- Build your own EKG

Vendor

- Understand how to develop, sell and deploy their products to maximize applicability and success

Software

*Architect, Vendor,
Academic*

- Access reusable software curated for EKG purposes (may be hosted externally)

Vendor, Academic

- Position existing products or components to increase uptake

Ontologies

Modeler

- Assess ontologies curated as reusable for EKG use cases
- Apply tooling to automate ontology development and measurement

Academic

- Access a set of ontologies for research, analysis, or extension

Datasets

Analyst

- Access reusable RDF resources curated for EKG purposes (may be hosted externally)

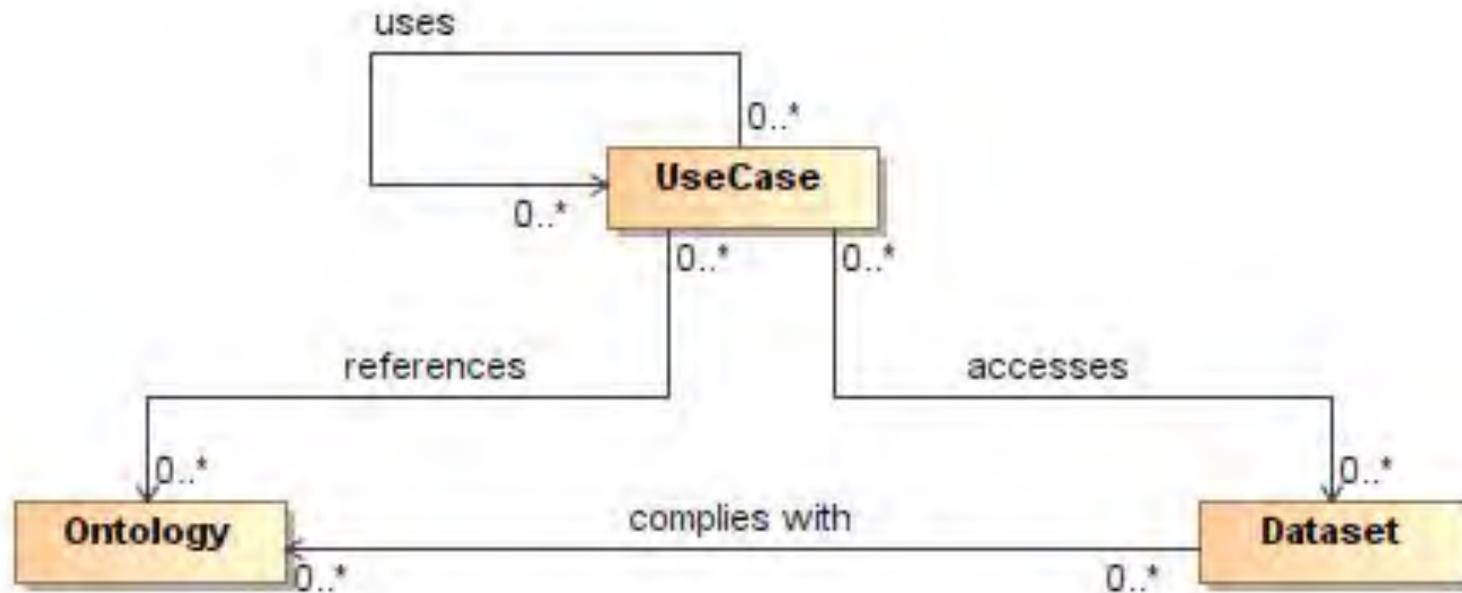
Analyst, Vendor

- Make use of data resources for demonstration and experimentation

Member Directory & Services

Connect with people with the knowledge to help ; promote your knowledge and skills

Key Reusable EKG Elements



Use Case as a Reusable Component

Plan phase

Key attributes of a use case, owned by “the business”

1. Name & Business Description
2. Desired Business Outcomes
 - “Definition of Success”
3. Personas, Concepts & Terms
 - add examples i.e. input for test scenarios
4. Stories & Workflow
 - high-level but agreed, metrics based estimates
5. Tree structure
 - Break-down into -- existing or non-existing -- sub-use cases (some of which are reusable and some of which are specific to the parent use case)
 - Priority is to look up in the tree, not down, and define the longer term “strategic use cases” as well

Build phase

Add additional detail after reaching high-level agreement

6. Datasets
 - already mapped to Ontologies
7. Ontologies
 - mapped to Concepts (in Use Cases)
8. Test scenarios
9. Story-implementation details
 - all optional -- e.g. SPARQL, SHACL, SQL, workflow, audit, history, provenance, entitlements, caching policies, etc
10. Metrics
 - function point or story point-like metrics
 - lead / cycle time metrics
 - predict cost & delivery
 - based on metrics of previous use cases

Run phase

Additional detail added after deployment

11. Usage metrics, tied to “Definition of Success”

12. Life cycle, support SLA etc

Use Case Portal

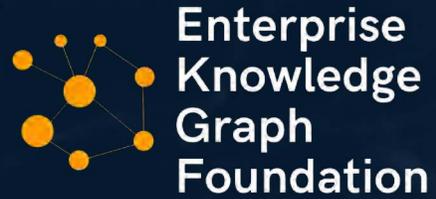
- Catalog of EKG-ready use cases
 - Focus on “strategic” use cases (the top level use cases)
 - Linked to “Business Outcomes”
- Use Case Trees
- Linked to ontologies and datasets
- Github based
 - Initially just as text albeit in RDF structure
- Leveraging “Github Actions” to:
 - Run basic “hygiene tests”
 - Run use case specific test scenarios with their test data in temporary triple stores (of any vendor)
 - Trigger tests of all dependent use cases
 - Generate various types of documentation
e.g. PDF, site, end user, EKG engineers, ontologist, quality
 - All software components to run this can also be deployed “on-prem”

Ontology Portal

- Catalog of EKG-ready ontologies
- Linked to use cases
- Tested per use case
 - No “boil the ocean” ontologies but ontologies that are thoroughly tested for production use cases
- All Git based
 - Could be any git repo
- Leveraging “Github Actions” to:
 - Run basic “hygiene tests”
 - Run use case specific test scenarios with their test data in temporary triple stores (of any vendor)
 - Trigger tests of all dependent use cases, ontologies & datasets
 - Generate various types of documentation
e.g. PDF, site, end user, ontologist, quality, use case specific info
- All software components to run this can also be deployed “on-prem”

Dataset Portal

- Catalog of EKG-ready datasets
 - Independent of use case i.e. “self describing datasets”
- Use cases are linked to datasets
 - Not the other way around
- Tested per use case
 - Since use cases define the real-world test scenarios
- Most are git based
 - Could be any git repo or facilities like data.world
- Leveraging “Github Actions” to:
 - Run basic “hygiene tests”
 - Run use case specific test scenarios with their test data in temporary triple stores (of any vendor)
 - Trigger tests of all dependent use cases, ontologies & datasets
 - Generate various types of documentation
e.g. PDF, site, end user, ontologist, quality, use case specific info
 - All software components to run this can also be deployed “on-prem”
- Goal is to make them automatically deployable



Thank You

Contact Us:

Web / Register: EKGF.org

Email: info@ekgf.org

Slack: ekgf.slack.com

LinkedIn Group: Enterprise Knowledge Graph Foundation