

# Enterprise Knowledge Graph / Maturity Model

---

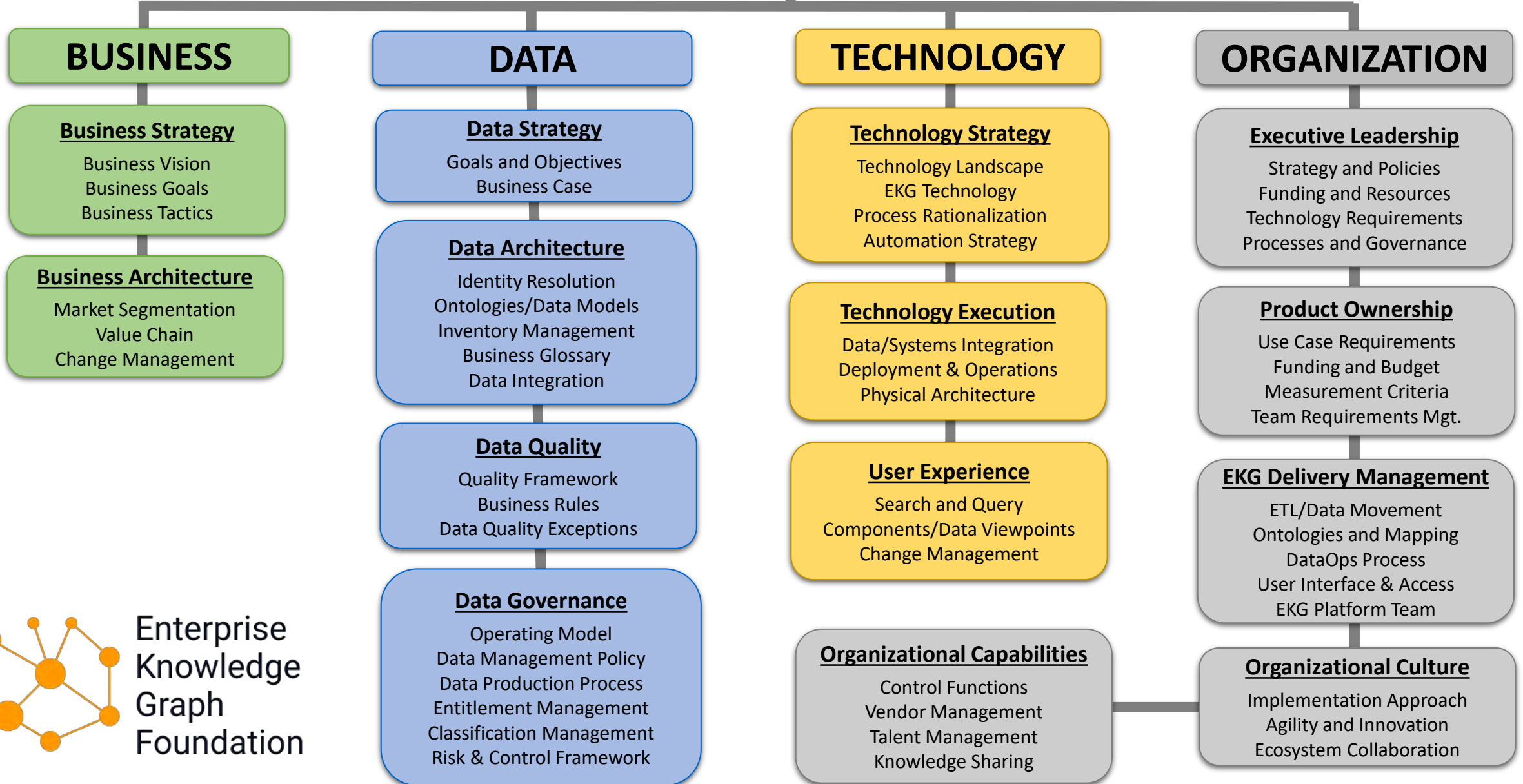
INAUGURAL MEETING • DECEMBER 17, 2020  
(STRUCTURE, COVERAGE AND DEVELOPMENT PROCESS)

# EKG/MM Perspective

---

1. Seed funding from [agnos.ai](https://agnos.ai) and Wizdom (contributed as an open-source initiative)
2. EKG/MM extends previous work on maturity modelling (CMMI, DCAM, Columbia)
3. Based on EKG Principles and Method
  - EKG Principles: identity, meaning, distributed sources, open world, self-describing, measurement, use case link, entitlement control, ecosystem, standards
  - EKG Method: strategic business focus → functional themes → use case tree → business concepts → data domains → data dependencies → core data attributes
4. GitHub as content repository and for issue management
5. Weekly working meetings (delivery goal: 2Q21)

# EKG Maturity Model Structure



# EKG/MM Maturity Expectation

## Initiate Journey

1

Implement foundational EKG components for minimum viable product. Establish the baseline architecture using isolated ontologies. Build initial team and processes. Project-based funding and specialist teams.

**Demonstrate Capabilities**

## Extensible Platform

2

Design ontologies for related (multiple) use cases. Create reusable architecture based on expanded design principles. Line of business alignment and funding. Center of Excellence and DataOps environment initiated.

**Parallel KG Activities**

## Enterprise Ready

3

Scalable and resilient EKG platform for mission-critical applications. Datapoint level authentication. Connected inventory linked to governance. Dedicated resources for design, build and support.

**EKG as Default Data Hub**

## Strategic Asset

4

EKG as strategic infrastructure for the organization and authoritative source for all glossaries. Supports application consolidation and cost-reduction. Leverages artificial intelligence and process automation.

**Operational EKG Utility**

## Operational Ecosystem

5

EKG is central to systems and business processes. Full integration into core business operations (including partners). Elimination of data silos. IaC architecture implemented. Inference and reasoning for advanced

**Continuous Development**

# Minimum Viable Product

## Level 1: EKG Initiation

The domain of pilots and internal POCs. The focus is usually on targeted use cases constructed using isolated ontologies. The champions are visionaries who have assembled a specialist team for implementation. Funding is likely to be project-based and designed to demonstrate capabilities.

<b>Business</b>	Stakeholders recognize the business liabilities from silos and data incongruence. Internal champion is seeking to solve strategic use cases, supports innovation and is willing to take on the disruption challenge.
<b>Data</b>	Core data management capabilities (operating model, inventory, data architecture, business glossary, pipeline management, etc.) are being performed. Specific use cases are being implemented with specialist teams for the pilot initiative.
<b>Technology</b>	Technology strategy is focused on experimentation and innovation. Manual data transformation and targeted ETL is underway for the pilot. Limited infrastructure and dedicated efforts to build initial knowledge graph components.
<b>Organization</b>	Champions are internal visionaries who have assembled a specialist team for implementation. The pilot is sanctioned and funded. Knowledge acceleration is being addressed. Overall organizational support is emerging.

# Reusable Components

## Level 2: Extensible Platform

The domain of parallel knowledge graph activities. The organization is creating reusable architecture based on expanded design principles. The EKG Center of Excellence is created. Funding is likely to be at the line-of-business level.

<b>Business</b>	Stakeholders adopt a “data centric” mindset focused on strategic business value. Management elevates the knowledge graph as an organizational and funding priority.
<b>Data</b>	Critical data elements are prioritized in the ontology. Approach to identity and meaning resolution is established. Use case trees are defined and modeled to capture shared data relationships. The knowledge graph is becoming the central point for integration.
<b>Technology</b>	Reusable architecture based on expanded design principles. Core software development design approaches are being established and incorporated into strategy. CTO focuses on extending pilot initiatives for additional leverage.
<b>Organization</b>	Operating model of collaboration is implemented to support the knowledge graph. The Center of Excellence and DataOps environment is initiated. Budget and implementation strategy are based on agile and synchronized with the use case tree methodology.

# Level 3: Enterprise Ready

The establishment of a scalable and resilient platform for business-critical applications. Resources for the design and build of operational systems are defined and coordinated. The knowledge graph is now an “enterprise” knowledge graph that serves as the data hub for the organization. Ownership, governance and funding are managed at the enterprise level.

## Default Data Hub

<b>Business</b>	Strong collaboration between business, data and technology to prioritize strategic (mission-critical) use cases.
<b>Data</b>	Inventory is embedded into the EKG and linked to governance. Data is expressed as formal ontologies, onboarded into the EKG and searchable. Data flows are defined and modeled. The EKG is the authoritative source for data.
<b>Technology</b>	Commitment to the EKG as the strategic infrastructure for the organization. IaC and continuous deployment adopted and implemented. Cloud architecture defined for elasticity. Datapoint security and authentication processes implemented.
<b>Organization</b>	The EKG is recognized as a core service for the enterprise. Enterprise-wide ownership and funding processes are operational. The EKG Center of Excellence is a stand-alone production department.

## Levels 4 and 5: Moving into the Future



<b>Level 4: Strategic Asset</b>	The EKG is understood as strategic infrastructure for the organization and the authoritative source for all data. It supports applications consolidation, AI and process automation. Strategic funding is based on the vision of executive management and fully embraced by the Board of Directors.
<b>Level 5: Operational Ecosystem</b>	The EKG is central to systems and business processes. It has been fully integrated into both internal operations and external supply chain partners. Workflows and approval steps are fully automated. Entitlements and access rights are controlled by the EKG. Inference and reasoning capabilities are used for advanced AI applications.



# EKG Maturity Model Sample

Standard  
Numbering

## B. DATA Pillar B.2 Data Architecture | B.2.1: Identity Resolution

**Description:** Identity resolution is a process of combining multiple identifiers across devices, spreadsheets, repositories and platforms into a cohesive profile. The process includes searching across disparate data sets and analyzing content to find (and resolve) matches based on available data records and attributes. Identity resolution is complicated by distinctions in both structure and meaning because various information systems can vary in quality, completeness, format and nomenclature.

**EKG Rationale:** Without standards for describing identity attributes, the process of resolution can be both time-consuming and risky. The knowledge graph uses formal descriptions (ontologies) of the terms in any domain as well as the relationship between terms. The use of ontologies and graph capability not only allows you to resolve multiple identifiers into a harmonized profile – it allows you to create associations across multiple identities with one master ID. The knowledge graph becomes the Rosetta stone for identity resolution.

Business  
Description

EKG  
Rationale  
and Value  
Proposition

## Core Data Management Questions

- **B.2.1.1:** Is there a standard process for identifying (inventory) and resolving (cross-referencing) identification schemes
- **B.2.1.2:** Does the firm maintain an inventory of identifiers for all strategic data assets
- **B.2.1.3:** Is the identification and meaning of data assets linked to the authorized systems of record
- **B.2.1.4:** Are versioning and time stamp management consistent across all data assets
- **B.2.1.5:** Is the method of generating and controlling identifiers aligned with governance process (i.e., defined by policy, verified, implemented and audited)

Core Questions  
(regardless of  
implementation  
approach)

## Common Principles for KG/IRI

1. All URLs have no meaning (opaque and meaningless)
2. Goal is to maximize proliferation of IRIs across the organizational ecosystem
3. IRIs are unique and permanent (no reuse and no elimination)
4. IRI naming conventions (domain and host names) must be sustainable forever (i.e., don't use company name because it might change if merger)

Pre-Scoring  
Orientation  
(based on EKG  
Principles)

# EKG/MM Scoring Sample (identity)

## Level 1 – EKG Initiation

- Describe and agree on the identifying properties of the entity (primary key attributes) and how IRIs will be constructed
- **[registration]** establish host and domain names for URLs/IRIs (per KG platform deployments)
- **[policy]** align approach with the KG IRI policy (add establish policy to policy section) - i.e., must have one, structure of IRI/hashing standard, maintain in inventory – goal is as short as possible (short, opaque, meaningless) – must decide and agree on standard (format and encoding)
- **[mapping]** all onboarding in-scope ETL pipelines must implement policy (generate the right IRIs). Ensure that all URLs and IRIs look the same (create and implement rules/formula for conversion)
- **[resolution]** define and implement process for resolving identical objects from multiple datasets into merged IRI (fully supporting multiple IRIs per instance)

Capability  
Requirements

Link to  
Governance

## Level 2 – Extensible Platform

- **[strategy]** develop strategy for all data (no longer compartmentalized) – strategy covers all domains across the ecosystem
- **[creation]** implement the model the IRI properties (per concept) into the knowledge graph
- Mapping of content to IRI is standardized using software routines

- Other systems also generate KG IRIs (extend mapping of IRIs to other systems and applications within the organization to prepare for enterprise conversion)
- Implement ability to look-up existing IRIs (reference hub) to allow other systems to link/integrate to KG (prerequisite for enterprise-wide KG)
- **[policy]** for the proliferation of IRIs (if the goal is a standard for creating for identification - is the standard - endorsed, mandated, and implemented) – includes policy for domain names (unchangeable) and commitment from executive management to register the top-level domain
- **[policy]** for synchronization of IRI with policy for objects (ontologies)
- **[policy]** for creation and control (only authorized systems are allowed) and policy for who is eligible to create IRIs (i.e., a certificate is needed to assign; only certain processes are eligible for certificates)

## Level 3 – Enterprise Ready

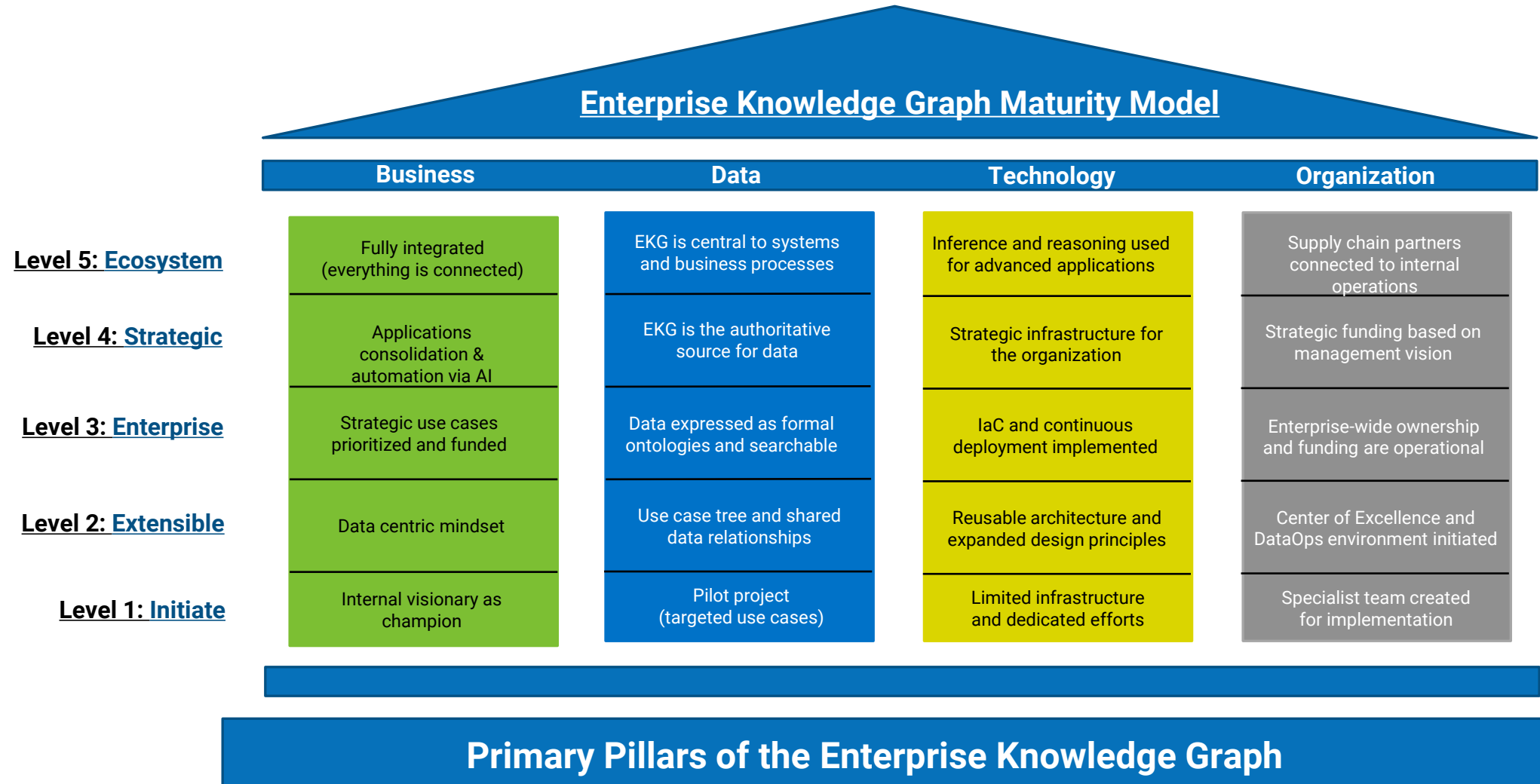
- Security procedures (i.e., specific crypto-security certificates) implemented to ensure control over IRI assignment process
- Implement standard IRI class hierarchy structure to enable multi-location linkage (shift from IRI “flavor” to IRN “flavor”)
- All incoming data flows are using the look-up service to convert code “strings” to the IRI “thing” (model driven/automated)

Implementatio  
n Advice

Policy  
Recommendations

Expectation  
Management

# EKG/MM Summary View



# EKG/MM Contribution Process

## Contribution Process (version 1.0 goal: 2Q21)

1. Research: weekly working group meetings and interactive discussions \*
  - EKG/MM Structure: January 14, 21  
(*open discussion about structure, scoring and operating model*)
  - Working Groups Development Process  
(*Data, Technology, Organization, Business*)
2. Sign-Up: [www.ekgf.org](http://www.ekgf.org) (click on members area)
3. Draft: EKG Foundation (GitHub + LaTeX for collaborative publishing)
4. GitHub <https://github.com/EKGF/ekg-mm/> (click on: “ekg-mm”, click on: “sections”)  
<https://github.com/EKGF/ekg-mm/projects>
5. Publish: EKG Foundation Website (<https://www.ekgf.org/maturitymodel>)



## Outside In

EKGF Website

GitHub

- Repository

.PDF files

Local Editors

## Inside Out

- Continuous Integration
  - Building and Testing
    - GitHub
      - Issues
      - Pull Requests
      - Publishing

Ways to engage

# Just Curious?

---

[www.ekgf.org](http://www.ekgf.org)

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

You will see a .pdf that you can read or download. It is complete, including scoring which will soon be reserved for paying EKGf members only

Want to work with us now?

# Want to See Everything (and you are an EKGf member)

[www.ekgf.org](http://www.ekgf.org) (click on members area)

- You will be asked to log in or to join if you have not previously joined
- You will follow the usual process and have a verification email sent to the email that you chose for your EKGf account
- When you log in, you will be able to download the complete MM

Want to contribute?

# Want to Contribute (start/join a work group)

BUSINESS

DATA

TECHNOLOGY

ORGANIZATION

- Membership is volunteer (co-chairs plus others)
- Meetings will be on a regular schedule
- Sticky notes or GitHub to:
  - Review content
  - Collect data
  - Document issues
  - Review Issues
- GitHub to:
  - Write Pull Requests
  - Approve Pull Requests
  - Automatically Publish content

The System of  
Record is GitHub



# GitHub for Content Management

---

Log into <https://github.com/EKGF/ekg-mm/>

We are developing a process to use Automated Kanban with reviews

It is *kind of* started <https://github.com/EKGF/ekg-mm/projects>

WE SINCERELY WELCOME  
EXPERIENCED HELP

Free form creativity is allowed (*prior to GitHub*)

Review and work with what we have now  
Tell us what we can do better together

Dennis.Wisnosky@ekgf.org

Text +1-630-240-6910

# Join the Enterprise Knowledge Graph Foundation

## Web Portals

### Best Practice

Maturity Model

Governance

Principles

EKG Method

DataOps

### Use Cases

Use Case Tree

Personas

UC Repository

### Quality

Ontology & Data

Tools

## Business Case

EKG Value Proposition

### Advocacy

Regulatory

Events

Publish

ROI Calculators

Case Studies

Benchmarking

Data Sets

Concepts

Taxonomy

Terms

## Membership

Content Repository

Webinars

Community

Directory/Intros

Education

Job Board

External Liaison

Influence

Vote/Prioritize

Leadership Roles

Working Groups

Join: <https://www.ekgf.org/membership>  
Principles: <https://www.ekgf.org/principles>  
Maturity: <https://www.ekgf.org/maturitymodel>

# EKG Foundation Contacts

Michael Atkin (Director)

[Mike.Atkin@ekgf.org](mailto:Mike.Atkin@ekgf.org)

+1.240.602.8390

(calendar appointment)

<https://calendly.com/mike-atkin>

Dennis Wisnosky (Process)

[Dennis.Wisnosky@ekgf.org](mailto:Dennis.Wisnosky@ekgf.org)

+1.630.240.6910

Jeannine W. Stehlin (Operations)

[Jeannine.Wisnosky@ekgf.org](mailto:Jeannine.Wisnosky@ekgf.org)

+1.310.701.5095

Jacobus Geluk (EKG Architect)

[jacobus.geluk@ekgf.org](mailto:jacobus.geluk@ekgf.org)

+44 (0)7518 691758 *(text is best)*

(calendar appointment)

<https://calendly.com/jacobus-geluk>

Pete Rivett (Ontologist)

[Pete.Rivett@ekgf.org](mailto:Pete.Rivett@ekgf.org)

+1.701.566.9534

EKG Foundation Website

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