Enterprise Knowledge Graph Foundation

EKG/MM

Overview of the Maturity Model for the Enterprise Knowledge Graph

February 21, 2022, Version 1.0.1.jacobus

Contents

Part I

Intro

Chapter 1

Preface

The Enterprise Knowledge Graph Foundation (EKGF) was chartered in early 2020. Of the many ideas discussed for initial projects, the founders agreed that writing and publishing a Maturity Model for the Enterprise Knowledge Graph (EKG) should be the first effort of the membership.

The idea of representing the Maturity Model for the Enterprise Knowledge Graph (EKG/MM) as four pillars titled Business, Organization, Data, and Technology, was presented in a June 2020 kickoff webinar.

This concept was accepted. Weekly pillar Zoom working sessions began shortly thereafter. These ongoing workgroups debate the contents toward the achievement of consensus in each pillar. Once a week the team leaders meet to synchronize content.

Our workgroups are lively and conversational. They are filled with people who are knowledge experts in their fields. They are often quite vocal in their points of view — always striving to make their work better and better and better. However, an Enterprise Knowledge Graph must, by definition, represent the views of the collective, not the views of individuals. Therefore, collectively, we present Maturity Model Release Version 1.0. Some may call it a draft version. Readers may find portions unfinished, and some sections say, "We welcome your input here." This input will go into the EKGF Continuous Improvement Process.

This is where you come in. Join our workgroups and participate in working towards Release Version 2.0.

The Enterprise Knowledge Graph Foundation team
info@ekgf.org

Chapter 2

Executive Summary

The EKG/MM is the industry-standard definition of the capabilities required for an EKG and of the generic capabilities in any given organization that are affected by EKG.

It establishes standard criteria for measuring progress and sets out the practical questions that all involved stakeholders ask to ensure trust, confidence and usage flexibility of data. Each capability area provides a business summary denoting its importance, a definition of the added value from semantic standards and scoring criteria based on five levels of defined maturity.

The EKG/MM is a capability model designed to promulgate best practices across the knowledge graph community. It covers essential capabilities as well as standard evaluation criteria required for the design, implementation, and maintenance of an EKG.

Chapter 3

Structure

The Maturity Model for the Enterprise Knowledge Graph (EKG/MM) is a collection of capabilities structured in a relatively simple 4x5 matrix of *capability pillars* and *maturity levels*.

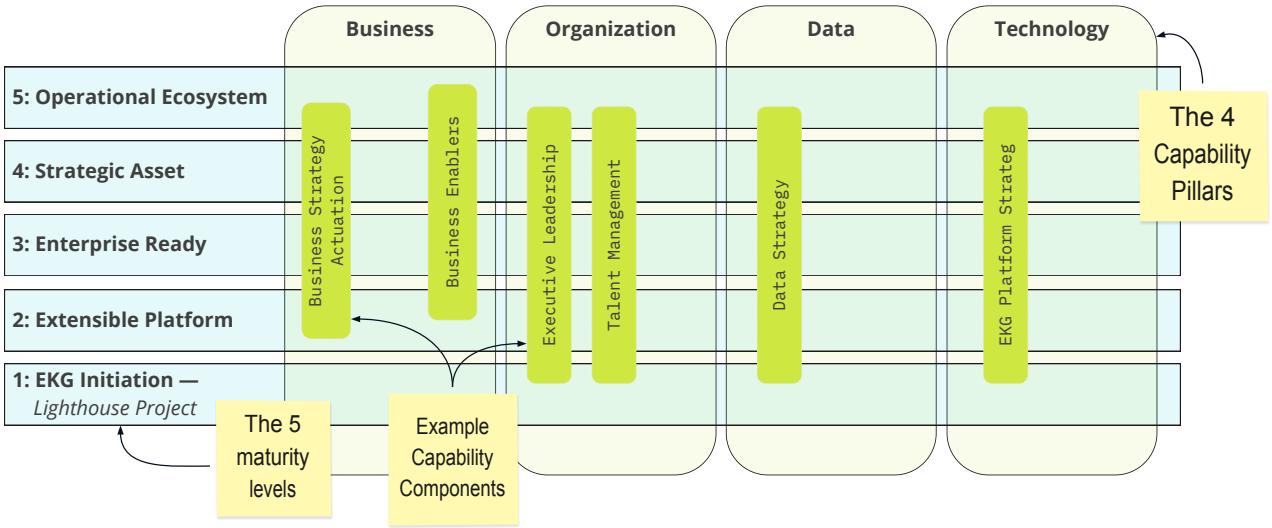


Figure 3.1: EKG/MM Structure

3.1 The Four Capability Pillars

All capabilities that we measure and assess in the EKG/MM[1] are categorized in four major categories — also called "pillars" — that match the four primary audiences that are involved:

The Business (Strategy) is leading. In the Business Pillar we discuss (and assess) our capabilities to formulate what our business identity is, our strategy actuation and elaboration and how that relates to business enablers and to the EKG. The other three pillars are facilitating but in a "Data-centric" and ultimately fully model-driven organization, business strategy gets translated to data strategy Data Strategy and technology strategy gets derived from both. The organization pillar discusses all capabilities that need to be in place to actually get all these strategies executed, with the right planning, budgets, processes and workforce.

These four pillars roughly correspond with the four primary audiences that we are addressing:

1. The audience of people on the business-side
2. The audience of people that are project managers, trainers or HR, legal, compliance and finance departments
3. The audience of people in the data-management and data-governance departments
4. The audience of technologists

[1] See [EKG/MM]

3.1.1 Business Pillar

All Business-side capabilities including Business Strategy Actuation, Business Model Elaboration, Business Enablers, Alignment, Operating Model and others. Does not include all the actual Business Capabilities that an enterprise may have themselves.

Addresses the audience of personas on the business-side of an enterprise, C-level, Line of Business (LOB) execs, corporate planners, business architects, product owners, management consultants and so forth.

3.1.2 Organization Pillar

All relevant Organizational Capabilities including Executive Leadership, Product Ownership, EKG Delivery Management, Organizational Culture, and Organizational Capabilities.

Addresses the audience of people that are neither business, data nor tech such as financial execs and experts, risk execs and experts, program/portfolio/project managers, HR execs and experts and so forth.

3.1.3 Data Pillar

All Data (Management) capabilities including Data Strategy.

Addresses the audience of people in the data-management and data-governance departments.

3.1.4 Technology Pillar

All Technology Capabilities including Technology Strategy, Technology Execution, and User Interface.

Addresses the audience of technologists, technical architects, developers, infrastructure execs and experts, security execs and experts etc.

3.2 The five maturity levels

This document will refer — for each capability — to the five maturity levels.

3.2.1 Level 1: EKG Initiation — Lighthouse Project

The domain of internal Proofs of Concept (PoCs), pilots or — ideally — "Lighthouse projects". The focus is usually on specific targeted baseline use cases constructed using isolated ontologies. The champions are visionaries who have assembled a specialist team for implementation — possibly the start of the CoE for the EKG. Funding is likely to be project-based and designed to demonstrate capabilities.

Business Stakeholders recognize business opportunities in scaling and amplifying capabilities through EKGs. The first internal champion is seeking to socialize strategic business cases, supports innovation, and is willing to take on the disruption challenge.

Organization 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.

Data Core data management capabilities (operating model, inventory, data architecture, vocabularies, business glossary or terminology, pipeline management, etc.) are being performed. Specific use cases are being implemented with specialist teams for the pilot initiative.

Technology 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.

3.2.2 Level 2: Extensible Platform — Reusable Components

The domain of parallel knowledge graph activities, implementing multiple (related) use cases on the same platform. The organization is creating reusable architecture — i.e. "the EKG/Platform" — based on EKG/Principles. The CoE for the EKG is created. Funding is likely to be at the LOB level and starts to (partly) come from Business-as-Usual (BAU) budgets.

Business Stakeholders adopt a "knowledge-centric" mindset in their tactics to strengthen focus on strategic business value. Management elevates the knowledge graph as an organizational and funding priority.

Organization 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.

Data 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.

Technology Reusable architecture based on EKG/Principles. Core software development design approaches are being established and incorporated into strategy. CTO focuses on extending pilot initiatives for additional leverage.

3.2.3 Level 3: Enterprise Ready — Default Data Hub

The establishment of a secure, scalable and resilient EKG/Platform for business-critical strategic use cases. Resources for the design and build of operational systems are defined and coordinated. The knowledge graph is now really an Enterprise Knowledge Graph (EKG) that serves as the semantic data mesh for the organization. Ownership, governance, and funding are managed at the enterprise level and coordinated by the CoE for the EKG that oversees the full life-cycle of use cases from inception to deployment and beyond. Long-term "operate & optimize" processes are in place.

Business Strong collaboration between various business and support units to prioritize strategic business cases.

Organization The EKG is recognized as a core service for the enterprise. Enterprise-wide ownership and funding processes are operational. The CoE for the EKG is a stand-alone BAU department.

Data 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.

Technology Commitment to the EKG as the strategic infrastructure for the organization. Infrastructure as Code (IaC) and continuous deployment are adopted and implemented. Cloud architecture defined for elasticity. Data point security and authentication processes are implemented.

3.2.4 Level 4: Strategic Asset — Operational Utility

The EKG is understood as strategic infrastructure — as an operational utility — for the organization and the authoritative source for most data (except for data that originates from core legacy systems). It supports structural application rationalization, high-level and high-quality Artificial Intelligence (AI) that can take over many tasks from humans and process automation. Strategic funding is based on the vision of executive management and fully embraced by the Board of Directors. All core data management capabilities have been achieved.

3.2.5 Level 5: Operational Ecosystem — Continuous Improvement

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 use cases.

3.3 Capability Template

The general layout, in this document, of each capability looks as follows:

Capability

[The title of the capability]

Description

[Tagline, a short one sentence description of the Capability]

[Description of the capability]

Contribution to EKG

[How the capability contributes to the EKG]

[Including traceability to EKG/Principles where relevant]

Contribution to the Enterprise

[How the capability as part of the EKG contributes to the enterprise]

Dimensions

[General dimensions/questions being measured by the criteria at different Levels]

[The following are candidate general questions]

- Level of required skills available (e.g. trained people)
- Level of resources available (e.g. infrastructure, tools)
- Level of defined processes available
- Scope of applicability
- Questions specific to the capability

Levels 1 –5

[Description of level]

[Criteria, drawing on the Dimensions]

Chapter 4

Overview

Table 4.1: The taxonomy of pillars, components, capabilities, and the measurable abilities or summaries.

	Component		Capability	Measurable ability
A Business Pillar	A.1 Business Strategy Actuation	A.1.1	Business Vision	Communicating the essence of what a business wants to achieve.
		A.1.2	Business Goals	Specifying the outcomes along the journey to achieve the Business Vision.
		A.1.3	Business Tactics	Ensuring that change capabilities are focused on business value realization.
	A.2 Business Model Elaboration	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.2.3	Change Management	Manage operations change to a new state of sustainable value.
	A.3 Business Enablers	A.3.1	Operating Model	The desired level of business process integration & standardization.
		A.3.2	Performance Management	Measure business processes to manage performance.
		A.3.3	Risk Management	Identify risk to mitigate or accept, with reaction plans.
		A.3.4	Supply Chain Management	Managed flow of goods, services & information between businesses.
		A.3.5	Capability Map	Defining what a business does and which outcomes it wants to achieve.
B Organization Pillar	B.1 Executive Leadership	B.1.1	Organizational Strategy	Organizational mandate and strategic support for the EKG.
		B.1.2	Organizational Policies	EKG alignment with the views and values of the organization.
		B.1.3	Funding & Resources	Funding process and resource allocation for the EKG.
		B.1.4	Technology Requirements	CIO support for EKG (where EKG fits in the overall technology stack).
		B.1.5	Organizational Processes	Alignment of EKG with business architecture.
		B.1.6	Organizational Governance	Process for aligning EKG with organizational accountabilities.
	B.2 Product Ownership	B.2.1	Use Case Requirements	Identification/prioritization of EKG use case requirements.
		B.2.2	Funding & Budget	Annual and multi-year funding process.
		B.2.3	Delivery Process	Management of statements of work for business outcomes.
		B.2.4	Measurement Criteria	Business metrics and criteria for evaluating the EKG.
		B.2.5	Team Requirements Management	Identification and management of skill sets needed for the EKG.
	B.3 EKG Delivery Management	B.3.1	ETL/Data Movement	Process for extracting/transforming data from source-to-EKG.
		B.3.2	Ontologies & Mapping	Approach to concept modeling and transformation.
		B.3.3	DataOps Process	Rules and standards for building, testing, and maintaining the EKG.
		B.3.4	User Interface & Access	Methodology for facilitating end-user access to data.
		B.3.5	EKG/Platform Team	Requirements and accountabilities for managing the EKG/Platform.
	B.4 Organizational Culture	B.4.1	Implementation Approach	Build versus buy decisions.
		B.4.2	Agility & Innovation	Empowerment, adaptability and failure tolerance.
		B.4.3	Ecosystem Collaboration	Alignment with technology, business, and operations.
	B.5 Organizational Capabilities	B.5.1	Control Functions	Coordination with privacy, legal, security, records management.
		B.5.2	Vendor Management	Acquisition and licensing of vendor products.
		B.5.3	Talent Management	Skills acquisition, retention and training.
		B.5.4	Knowledge Sharing	Ability of the organization to facilitate knowledge collaboration.
C Data Pillar	C.1 Data Strategy	C.1.1	Goals & Objectives	EKG as the semantic data mesh for the organization.
		C.1.2	Business Case	Business rationale, justification and Return On Investment (ROI) of the EKG.
	C.2 Data Architecture	C.2.1	Ontologies	The *models* that provide the structure and meaning for the information.
		C.2.2	Business Vocabularies	Terms used to refer to the information by different business communities.
		C.2.3	Datasets	Coherent collections of information made available for access.
		C.2.4	Data Mapping	Unifying or relating different information concepts.
	C.3 Data Quality	C.3.1	Data Quality Business Rules	Strategy and approach for managing data quality.
	C.4 Data Governance	C.4.1	Data-management Operating model	Structure and approach for the EKG Center of Excellence.
		C.4.2	Data-management Policy	Policies, procedures, and standards for managing the data lifecycle.
		C.4.3	Classification Management	Management of the data production and manufacturing process.
		C.4.4	Risk & Control Environment	EKG access control mechanisms.
D Technology Pillar	D.1 Technology Strategy	D.1.1	Technology Architecture	Plans for physical infrastructure, applications, and automation.
	D.2 Technology Execution	D.2.1	Data Integration	Ability to source and integrate data (consolidate silos).
		D.2.2	EKG Operations	Deployment and maintenance requirements for the EKG.
		D.2.3	Logging / Systems monitoring	Physical infrastructure and composite view of systems/applications.
		D.2.4	Business Continuity	Continued business operations during unplanned EKG/Platform disruption
		D.2.5	Data Quality Implementation	Technical implementation of data quality framework
		D.2.6	Analytics	Technical implementation of analytics facilities
		D.2.7	Inferencing Rules	Inferring new facts from existing facts by means of ontologies
		D.2.8	Entitlement Enforcement	Enforcing fine-grained context-specific entitlement policies
		D.2.9	Data Transformation Execution	Governed automated orchestration of data transformation
		D.2.10	Identity Resolution	Standards for ensuring content is identified and resolvable.
		D.2.11	Knowledge Graph Storage and Retrieval	Semantic data storage managed by the EKG/Platform
		D.2.12	Knowledge Graph Federation	Model-driven federated interoperable and distributed data platforms
		D.2.13	Knowledge Graph Virtualization	Model-driven views on data of any source leaving it in-place
	D.3 User Interface	D.3.1	Search	Standards and tools for finding relevant data across the enterprise.
		D.3.2	Collaboration	...
		D.3.3	Multilingual Support	...
		D.3.4	Content Editing	...
		D.3.5	Reporting	...
		D.3.6	Visualization	...

Part A

Business Pillar

Introduction & Background

The modern economy is often characterized as an Information Economy or more specifically, Knowledge Economy, emphasizing the importance of knowledge as the key productivity lever and intellectual capital as the key differentiator of enterprises. Widespread digitalization over the last two decades has accelerated the scale, reach and impact of businesses in today's knowledge economy. Seven of the top ten firms globally by market capitalization at the time of this writing (Apple, Amazon, Microsoft, Alphabet, Facebook, Tencent, Alibaba), are representative of this characterization. This underscores the broad expectation that future competitiveness is increasingly focused on embedding and capitalizing knowledge in the design, delivery and eventual effectiveness of products, services and activities in business models of Enterprises.

What does a knowledge-based differentiated focus in products and services mean for enterprises? Prevalent characterization of some business models offers some clue:

(a) Embedded Finance — the emphasis that convenience of financial products is increasingly determined by integrating them seamlessly into non-financial applications and processes. Embedded Finance offers a new, very large addressable market opportunity worth over $7 trillion in ten years' time, twice the combined value of the world's top 30 banks today.[Tor]

(b) Subscription Services — the emphasis that experience of products and services is increasingly determined not from broad contours of use over time but by specific contexts of use at any given time. 78% of international adults in 2021 have subscription services (71% in 2018). And 75% believe that in the future, people will subscribe to more services and own less physical "stuff". [Kalim]

(c) Platform Services — the emphasis that utility of services is increasingly determined not from clarity of individual service but through convenience of search and similarity of use across a range of related products and services. More than 30% of global economic activity — some $60 trillion — could be mediated by digital platforms by 2025. [Hirt]

In the following sections, we explore a few illustrative capabilities that enable enterprises to leverage and embed knowledge in their propositions, products and services, planning and activities.

2019 — When we exceeded 1 billion Knowledge Workers

...Where we're going is a future of work with knowledge... being created at the pace only possible with one billion workers. What does empowering its knowledge workers and embedding knowledge in products and services mean for enterprises?

The Business Case for Explorative Execution

Business performance is traditionally measured with metrics emphasising financial performance — typically, aspects of revenue, profit, growth and risk. Mechanisms such as Balanced Scorecards help enterprises align strategy execution to their business performance, emphasising the causal chains coursing through organizational capability, operations and customer propositions, eventually landing into financial performance as an outcome. More recently, there has been an emphasis on key stakeholder metrics such as ESG adding further dimensions to responsible business performance.

The paradigm of business performance has largely hinged on organizations directly influencing their customer engagements with their products and services. Indeed, the prevalent use of the term "customer" is ingrained as a counterpart to the organizations' products and services. Dramatic advances in the digital economy, especially in the area of mobile and sensor technology over the last 15 years, have highlighted the perception of value through insights into consumption/use. In other words, the presumed value in organizations' products and services is increasingly validated through actual insights into consumption/use by consumers.

This shift in perspective of an organizations' products and services as patterns (based on insights) of consumer use is where movements such as design thinking, customer journey maps etc are increasingly considered as important tools for

developing and validating propositions. The ”aha” and ”viola” moments from such exercises require further translation into realizing the propositions, which is where practices around agility, DevOps etc are increasingly mainstream.

All these themes and mechanisms of strategy execution through communication, product/service propositions, development/deployment and its performance management largely depend on a human capital based on participation, expertise and intuition. Organizations wishing to amplify their human capital face a few key dichotomies to be addressed:

- **Customer vs Consumer**: Mindset and capability shift from “Customer market share” to “share of (end) consumer value”
- **Enterprise vs Ecosystem**: Mindset and capability shift from “Proprietary control of customer experiences” to “open collaboration on consumer experiences”
- **Human Machine Continuum**: Mindset shift from human-scaled to machine-scaled (intelligent and autonomous) business capability

Each of the above dichotomies highlights the limitations enterprises face today in extending business capability beyond their organizational boundaries, be it in sourcing information beyond their own channels or collaborating on shared experiences.

Addressing these challenges requires enterprises to adopt an exploratory style in executing their strategy[1], which specifically focuses on sensing and understanding the consumer ecosystem better before taking decisions and acting (to be aligned to hyper-personalization, both in retail and business consumer scenarios)

To put this “Explorative Execution” into perspective, a useful mental model is Sense, Understand, Decide, Act (SUDA) (a slight variant of John Boyd’s OODA loop), specifically focused on responding/acting in a dynamic evolving environment. The choice of the mental model is to bring out its contrast with the imperative style mental model of Plan, Do, Check, Act (PDCA), usually used for acting in an environment of more certainty and predictability.

SENSE Accessing actual contents of need and use by consumers easily overwhelms most organizations for a variety of reasons, chief amongst which is the sheer volume of context-unique interactions which must be checked for links to an organization’s business interests and intents. Most of this is practiced subjectively and accorded to human expertise and intuition.

UNDERSTAND Unlike the carefully curated lens of its products and services, inputs of interest for an organization are generally ambiguous requiring significant efforts to disambiguate and synthesize. (John Boyd’s Orientation isn’t just a state you’re in; it’s a process. You’re always orienting.)

DECIDE Disambiguated information still isn’t directly actionable as there is a significant element of hypothesis embedded within. This is especially challenging with high prevalence of conflicting hypotheses in many cases.

ACT Filtered sensing, hypothesis-based understanding and scenario-driven decision-making finally create an actionable context, which the organizational operations can execute.

A deeper analysis of the SUDA loop of “Explorative Execution” brings into focus the following requirements of capability:

Identity how can sensing be practiced in a subject domain of multiple identities? How can a shift from form-based to a function-based identity be enabled?

Self-describing how can understanding be practiced in a subject domain with multiple descriptions? How can a shift from structure-based description to a behavior-based description be enabled?

Open world ambiguity how can decision-making be practiced in scenarios which require constant revalidation and conflict resolution, as new facts become known?

An EKG is a key organizational capability to enable Explorative Execution using a SUDA mental model, through the following core features:

SENSE
- Co-existence of multiple identifications based on context and interpretation instead of the prevalent content based identification
- How do retail consumers and business consumers use products and services?

UNDERSTAND
- Emphasis on meaning through association rather than static descriptions
- Enhanced collaboration with ecosystem partners through standards on context sharing rather than lengthy negotiations on content specifics
- How do the products and services address needs of consumers?

DECIDE
- Significant scale in knowledge driven operations through autonomous reasoning
- Leverage of existing (diverse and federated) content sources instead of separate curated content repositories
- How can products and services add more value through effective use?

[1] See ‘What is Strategy?’ on page 84

- What changes can be made to make products and services more relevant for (retail and business) consumers?

Business Identity

Business Identity is a statement of a company's unique innovation, service, and/or product based value proposition. A Business Identity, published to the world via business identifier unique Enterprise Knowledge Graphs, may transparently use a data infrastructure that can support many distinct business activities.

An enterprise could have multiple Business Identities based on the value propositions underlying its products and services — it is imperative that each must be perceived (as having differentiators) and consumed (as unique offerings) by its customers and stakeholders.

- The **Sense** imperative is addressed by using the Business Identity as an "attractor" amidst the voluminous stream of interactions. This is especially important in a privacy conscious environment where interest in and intent of using information is clearly understood and agreed upon upfront.
- The **Understand** imperative is addressed by using the Business Identity as a common link between the complementary mechanisms of (a) fulfilment within an enterprise and (b) the collaboration contexts with the partners.
- The **Decide** imperative is addressed by using the Business Identity as an enterprise persona for scenario planning and evaluation- rather than viewing an enterprise as a structural entity.

Business Identity enables an (EKG powered) enterprise to constantly assess and align its value propositions to (a) the needs and use contexts of consumers and (b) the collaboration mechanisms with business partners for effective fulfilment.

Examples of prominent *brand statements*:

Apple Watch *It's the ultimate device for a healthy life. Apple Watch can do what your other devices can't because it's on your wrist. When you wear it, you get a fitness partner that measures all the ways you move, meaningful health insights, and a connection to the people and things you care about most. And it's always just a glance away.*

Netflix *Unlimited films, TV programmes and more. Watch anywhere. Cancel at any time. Netflix is a subscription-based streaming service that allows our members to watch TV shows and movies without commercials on an internet-connected device.*

A Business Identity can link brand statements such as the above with the wider contexts of an Enterprise's corporate values, vision etc and the specifics of the product/service features and integration with other products/services of the enterprise/wider ecosystem.

Aligning Business Identity with Activities in an Enterprise Enterprises use a variety of mechanisms today to organize "inventories" of their activities. The artifacts take a variety of forms ranging from free-form documentation of manuals, data and process models along different methodologies to data records and programming repositories. Business Capability Models, Business Canvas, Business Reference Architectures etc have been some of the outputs used to serve as organization-wide references of activities.

The objectives guiding such management are mostly around efficiency of categorizing, storing and retrieval, focused around "what is done" and "who does it". The key challenge in efficiency oriented approaches is that the perception can be quite subjective based on the needs of the departments or teams- which leads to a variety of approaches across the length and breadth of any enterprise. There are several implications of using such a variety of approaches — chief amongst which is the reconciliation effect expended every time a deviation has to be managed, exceptions have to be handled or a change introduced in any activity. Targeting efficiency of reconciliations through techniques such as focus-groups, standardization of documentation etc. is a typical response of most organizations.

Taking the intuitive notion of a Business Identity (as a conceptual anchor for an organization's engagement with its stakeholders) forward, enterprises can benefit from using a simple, organization-wide concept anchor focused around "why is it done" for its activities. This anchor need not be prescriptive to change the way an organization manages its activities but descriptive to enhance the ability to more effectively leverage information within the artifacts.

One candidate for such an intuitive anchor is the contract. In fact, a contract may initially take the form of a memorandum of understanding (MOU), an agreement outlined in a formal document that may later become a legal contract. If we take the Business Identity as a basis for viewing an organization's product and services, a contract can be seen as a specific engagement around which the activities of various departments and teams are organized. More specifically, a contract can be seen as a collection of related commitments, potentially spread over time, which provide the necessary context

for any activity in an organization. Interestingly, while it's well-understood that a contract is a definitive source of specifics around any agreement with stakeholders, it's currently referred to only in the extreme case of legal recourse in handling disagreements. Another perception of contracts is that they do not cover all collaboration contexts which guide activities in an organization.

Contracts have several features which help align the Business Identity with the activities within an enterprise — and the wider ecosystem, for that matter.

a Together with the legal system and societal norms, contracts provide a contextual basis for belief/trust between an enterprise and its stakeholders in the ecosystem

b Contracts are specific to the parties involved and provide an ideal contextual basis for (hyper)personalisation of products and services — through bounded simulation of contract performance

c Measures of performance in organizational activities, such as cost, operational efficiencies and innovation effectiveness can be aligned to contractual performance contexts

We now look at two key elements of organizational activity (enabled by EKG) aimed at orienting and realizing business strategy, through the lens of Business Identity and Contracts:

1. **Business Strategy Actuation** is a business identity oriented process that senses, understands and communicates business outcome-based assessment data to focus energy and resources on strategic objectives. Actuation measurement values are used to assess whether strategic objectives and associated business goals are being fulfilled.

   See component A.1 Business Strategy Actuation on page 19

2. **Business Model Elaboration** is the process of providing further detail of business strategy. It may include decision rationale that explains how the business identity data infrastructure and algorithms enable the how and why planned actions would create, deliver, and capture value.

   See component A.2 Business Model Elaboration on page 23

Chapter A.1

Business Strategy Actuation

Business Strategy Actuation is a business identity oriented process that senses, understands and communicates business outcome-based assessment data to focus energy and resources on strategic objectives. Actuation measurement values are used to assess whether strategic objectives and associated business goals are being fulfilled.

As a refinement of the above definition, **Business Strategy Actuation** is the process of identifying measurement values — e.g. speed, time — for planned object change activities — e.g. add, remove, move — that will provide an accurate way of assessing whether strategic objectives and associated business goals have been fulfilled.

Actuation is an inspired term — in common use, it denotes the action of causing a machine or device to start. In the business world, *Business Strategy* is about intent, and *Strategy Actuation* can be seen as activating that intent. The broad business imperatives of Sense, Understand, Decide, Act (SUDA) for a modern, forward-looking Enterprise, based on the "siren call" of its *Business Identity*, have been established in the introductory section on page 17.

This section suggests mechanisms wherein EKG can help Businesses in aligning their Strategy Actuation to the SUDA imperatives. While practices around Business Strategy vary across Enterprises — for historical reasons and management preferences — they broadly cover the following key elements of Business Intent and its cascade within the Enterprise:

Business Vision Communicating the essence of what a business wants to achieve.

For example:

- *Create a better everyday life for many people* — IKEA
- *Bring Inspiration and Innovation to every Athlete* — Nike
- *Create the most compelling car company of the 21st century by driving the world's transition to electric vehicles* — Tesla

Business Goals Specifying the outcomes along the journey to achieve the Business Vision.

Strategic Business Goals and their articulation have been human-facing primarily, with guidance and governance mechanisms taking forms such as periodic reviews, monitoring and reporting. While tools and workflows around SMART[1] principles have rendered more objectivity in the process, significant judgment is still required in assessing progress towards strategic goals. Such judgment involves a more holistic assessment of the progress taking a balanced view across objective measurements and evidences of impact.

Business Tactics Ensuring that change capabilities are focused on business value realization.

Business Tactics are the actions performed that ensure business change capabilities are focused on business value realization. To determine that added value is created by a particular business change, the key factors used for metric generation can be made explicit as in EKG.

Business Tactics, which specify the actions to focus and align business activities and pursuits to the realization of the Business Goals practices around agility in business change, as well as real-time focus on monitoring of key internal and external BAU metrics have enabled enterprises to make frequent tactical corrections in their activities. A significant component of such corrections is a judgment that relies on human intuition, bias and expertise, guided by associated business metrics.

[1] Specific, Measurable, Achievable, Realistic, Timely

To summarize, **Business Strategy Actuation** is a business identity oriented process that senses, understands and communicates business!outcome-based assessment data to focus energy and resources on strategic objectives. Actuation measurement values are used to assess whether strategic objectives and associated business goals are being fulfilled, such as:

- Speed of change — which propositions, and their contractual commitments, will benefit from an enhanced speed of change?
- Speed of awareness — which contractual commitments will benefit from an enhanced ability to assess and respond?
- Speed of understanding impacts — which commitments require a focus on demonstrating suitability through impact analysis?
- Speed of adapting to one or more changes — which commitments are most at risk on the speed of adapting to changes?
- Ability to identify errors and understand their impacts — how do exceptions relate to commitments and how does exception management affect commitments?
- Ability to measure error propagation and assess scale
- Ability to manifest change to respond to errors

The Business Strategy Actuation component has the following capabilities:

A.1.1 Business Vision — Communicating the essence of what a business wants to achieve.

A.1.2 Business Goals — Specifying the outcomes along the journey to achieve the Business Vision.

A.1.3 Business Tactics — Ensuring that change capabilities are focused on business value realization.

A.1.1 Business Vision

Description

Communicating the essence of what a business wants to achieve.

For example:

- *Create a better everyday life for many people* — IKEA
- *Bring Inspiration and Innovation to every Athlete* — Nike
- *Create the most compelling car company of the 21st century by driving the world's transition to electric vehicles* — Tesla

Contribution to the Enterprise

An enterprise's Vision and Mission have been mechanisms for easy retention and recall within its rank and file — for guidance and motivation in activities — as well as its public image. As communications and activities of enterprises become more digital, leveraging a mix of human and machine scale capabilities, *EKGs can help by enabling enterprises to align the Vision and Mission in the context of activities, through the Business Identity.* Importantly, the Business Identity anchored EKG can provide a contextual lens for articulation, validation and justification of Business Vision at any level of activity of an Enterprise — within as well as across its boundaries.

Contribution to the EKG

Having a clearly communicated Business Vision will help setting the direction for the discovery and selection of use cases for the EKG.

At higher levels of maturity (level 3 and up) the Business Vision itself would be a key use case for the EKG, allowing any other part of the EKG to be directly or indirectly relatable to the vision, goals, providing insights into which components of the organization (people, systems, products etc.) are supporting the vision and how.

A.1.2 Business Goals

Description

Specifying the outcomes along the journey to achieve the Business Vision.

Strategic Business Goals and their articulation have been human-facing primarily, with guidance and governance mechanisms taking forms such as periodic reviews, monitoring and reporting. While tools and workflows around SMART[2] principles have rendered more objectivity in the process, significant judgment is still required in assessing progress towards strategic goals. Such judgment involves a more holistic assessment of the progress taking a balanced view across objective measurements and evidences of impact.

Example objectives:[3]

- Stay Competitive.
  - Comprehensive, client-focused services.
  - Outstanding expertise and service quality.
- Increase agility reacting to or delivering on Risk & Regulatory requirements.
  - Comply with ethical, data privacy and security guidelines.
- Drastically decrease "time to market" of the implementation and realization of new business ideas from inception to production.
  - Offer more tailor-made services to each client.
- Silos are holding us back: lack of "360 degree views" and not having the full picture is a drag on competitiveness, no more (new) silos.

Contribution to the Enterprise

EKGs can help by enabling enterprises align expectations, evidences and business metrics both in achieving and assessing progress on Strategic Goals.

Contribution to the EKG

Having clearly communicated, actionable and measurable Business Goals will help setting the direction for the discovery and selection of use cases for the EKG.

At higher levels of maturity (level 3 and up), all Business Goals, at any level of granularity and scope, would end up being modelled and become part of the EKG, directly or indirectly linked to every activity in the organization, providing insights in cost, progress, effort and risks per stated goal.

A.1.3 Business Tactics

Description

Ensuring that change capabilities are focused on business value realization.

Business Tactics are the actions performed that ensure business change capabilities are focused on business value realization. To determine that added value is created by a particular business change, the key factors used for metric generation can be made explicit as in EKG.

Business Tactics, which specify the actions to focus and align business activities and pursuits to the realization of the Business Goals practices around agility in business change, as well as real-time focus on monitoring of key internal and external BAU metrics have enabled enterprises to make frequent tactical corrections in their activities. A significant component of such corrections is a judgment that relies on human intuition, bias and expertise, guided by associated business metrics.

[2] Specific, Measurable, Achievable, Realistic, Timely

[3] Taken from a public presentation at Enterprise Data World '21 of LGT Private Banking called "Enterprise Knowledge Graph in action at LGT"

Contribution to the Enterprise

EKGs can help by enabling enterprises to make faster tactical decisions, by providing an integrated view to consistently reason more effectively around metrics.

Contribution to the EKG

We welcome your input here.

Chapter A.2

Business Model Elaboration

Business Model Elaboration is the process of providing further detail of business strategy. It may include decision rationale that explains how the business identity data infrastructure and algorithms enable the how and why planned actions would create, deliver, and capture value.

For example:

- a *Business Canvas* models and elaborates value proposition, data infrastructure, customers, and finances as aligned activities with potential trade-offs.
- SUDA (Sense, Understand, Decide, Act), a lean form of PDCA (Plan, Do, Check, Act), elaborates how the future state data infrastructure supports the value stream of one or more value propositions.

The Business Model Elaboration component has the following capabilities:

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.2.3 Change Management — Manage operations change to a new state of sustainable value.

A.2.1 Market Segmentation

Description

Shared understanding of markets, products, and clients.

The capability "Market Segmentation" is the ability to structurally recognize, define, manage and use the company's market segments and drive decision-making around products, services, sales and marketing operations.

Market Segmentation is an important component in the elaboration of an Enterprise's Business Model, helping Enterprises to align their Product and Services fulfilment to the needs of their customers. Market Segmentation practices today involve an intensive analysis through a mix of surveys, heuristics and benchmarking. The availability of and access to voluminous information though a multitude of digital sources has put additional emphasis on the ability of enterprises to dynamically tweak their Market Segmentation.

> *"98% of firms agree that static segmentation is no longer adequate, and 100% of firms recognize the opportunity cost of not evolving to its more granular, precision-based, inheritor: Dynamic Predictive Segmentation."*
>
> *The Customer Moment With Dynamic Predictive Segmentation [18]*

Contribution to the Enterprise

Having the proper Market Segments defined will help with the selection, definition and prioritization of the right use cases for the EKG.

Contribution to the EKG

The EKG can bring all possible internal and external detail information about customers, products, services, competition, sales volumes, customer requirements and many other details together forming a holistic and realistic, almost real-time view of the position of the company in their markets.

A.2.2 Value Chain

Description

Alignment of business capabilities with EKG pathway.

The Value Chain capability is about having the Value Chain(s) defined, updated, used and managed.

Value chains have been a prominent way of organizing and linking activities in an Enterprise to create value for its customers, thorough provision of products and services. Digitalization and the increasing focus on consumer-centricity has pushed the imperatives of Value Chains to include more holistic collaborations in the business ecosystem — reshaping provision of services and access to untapped pools of value.

> *"Most global companies are now actively considering the ecosystem business model given its value-generation potential: growing the core business, expanding the network and portfolio, and generating revenues from new products and services. "*
>
> - McKinsey Digital[1]

Contribution to the Enterprise

In the EKG context, a "digital twin" of a company's value chain can be modelled where all the various components of the value chain are represented e.g. all details around logistics, supply chains, operations, services, marketing, sales and all support activities. At higher levels of maturity, all real-world details of these components are available as well and represent the reality accurately and in real-time.

Contribution to the EKG

Having the Value Chain(s) of an organization defined can help with the selection, definition and prioritization of the right use cases for the EKG.

A.2.3 Change Management

Description

Manage operations change to a new state of sustainable value.

> *"Change management — The business process that coordinates and monitors all changes to the business processes and applications operated by the business, as well as to its internal equipment, resources, operating systems, and procedures. The change management discipline is carried out in a way that minimizes the risk of problems that will affect the operating environment and service delivery to the users. "*
>
> - ASCM[2]

Change management guides the business for change required to react to problems or to proactively plan changes to mitigate risk, obtain a good result for divesting and/or acquiring new areas of business, product changes for shifts in customer needs, improve processes for efficiency and cost savings, stay ahead of competitors, take advantage of new technologies, etc.

Change management is the discipline to guide realization and sustainment of business strategy intent. Impact of changes in strategy should be modeled and serve as a basis for executing the change. Change can be managed for any aspect of the business to include: geographic and facility footprint, market expansion, product development, business organization and

[1] https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/how-do-companies-create-value%2Dfrom%2Ddigital%2Decosystems#:~:text=Most%20global%20companies%20are%20now,from%20new%20products%20and%20services.

[2] https://ASCM.org

staffing, new sales channels, process improvement, technology change, compliance to changes in regulations, and more.

Contribution to the Enterprise

The EKG can hold the model of the enterprise to identify potential areas for change to include: discover duplication across business silos, impact analysis of areas targeted to change, impact of regulatory change, impact of entering new markets, impact of scaling the business, impact from major incidents and issues, and can serve as a baseline for evaluating and executing merger/acquisition of new enterprises.

The EKG can hold the information on current state of the enterprise for reuse of information to start the project, ideal state of the enterprise, intended new state of the enterprise after the change is deployed, and updated for actual resulting business knowledge from deploying the change.

Contribution to the EKG

The EKG should be updated to reflect changes made to the enterprise as a result of change projects. EKG governance is a set of change processes that should follow a disciplined change approach to protect the integrity and value of the knowledge graph.

More background Change is happening everywhere in organizations. The digitalization of everything (IoT, edge computing, digital transform) has created a fluid context for change, and the need to decentralize decision-making to keep pace with exponential change and complexity requires non-linear thinking, planning, and execution.

The current approaches for change management are usually built on linear models which do not scale nor can they collate the facets of data — let alone semantic perspectives — that are overwhelming decision makers.

EKGs are the only way to build a scalable framework of understanding about change and capture relevant patterns to guide option analysis and impact analysis. The days of getting in a room and relying on meetings is quickly coming to an end. AI, Bayesian Analysis, Game Theory, multiple-criteria decision analysis (MCDA), and Heuristics can help, but only if we can guide them properly using EKGs.

> *"Ask not what Change Management can do for EKG, rather ask what EKGs can do for Change Management "*
>
> - Jeffrey Wallk[3]

[3] https://www.enablingvalue.com

Chapter A.3

Business Enablers

The Business Enablers component has the following capabilities:

A.3.1 Operating Model — The desired level of business process integration & standardization.

A.3.2 Performance Management — Measure business processes to manage performance.

A.3.3 Risk Management — Identify risk to mitigate or accept, with reaction plans.

A.3.4 Supply Chain Management — Managed flow of goods, services & information between businesses.

A.3.1 Operating Model

Description

The desired level of business process integration & standardization.

Operating Models are the foundations of execution, in support of an enterprise's strategy. Enterprises use operating models to organize and coordinate their activities in order to most effectively engage with their customers and business partners.

At its simplest, an Operating Model enables an Enterprise to realise its vision and strategy at the level of everyday execution — business operations.

An operating model — or lack of it — has a profound impact on how a company implements its business processes and organizes supporting technologies. Depending on an enterprise's Business Models, multiple operating models, at corporate and unit/divisional levels, could co-exist, in pursuit of the business strategy.

In order to define an operating model at any of these different levels, Enterprises need to answer the following two questions:

1. To what extent is the successful completion of one business unit's operations/transactions dependent on the availability, accuracy, and timeliness of other business unit's information?
2. To what extent does the enterprise benefit by having business units run their operations in the same way?

The first question determines integration requirements while the second one covers the standardization requirements of an Enterprise.

> *"80% of CEOs in one study claim to have transformations in place to make their businesses more digital; 87% expect to see a change in their operating models within three years."*
>
> Deloitte [Del20]

In today's modern dynamic economic environment, Enterprises have to contend with a variety of choices in the selection, ongoing evaluation and transformation of their operating models. Such choices are necessitated by a fast-changing competitive landscape and the accompanying need to constantly align the workforce capacity and capabilities to business tactics. Typical triggers of consideration include comparative metrics such as higher-than-peer Cost-to-Income Ratios in key product portfolios, costs per transaction, Net Promoter Score/CSI, time-to-market for new products et al.

While practices around specifics of Operating Models vary across Enterprises, we highlight below a few key focus areas in the setup and transformation of their operating models and in such a context, how EKGs can act as effective enablers:

Integration: Effective coordination across silos of activities with disjoint goals in any given external interaction (customer or stakeholder) In the Introductory section, we have introduced the idea of using Contracts as an intuitive anchor for coordinating activities in an enterprise, towards its engagements with customers and stakeholders. An Operating Model can leverage the context of ”why” and ”what” in the contracts and specify the ”who” and ”how” to align and link various activities within a business unit, across business units and across enterprise boundaries. Many enterprises are using the concept of ”User/Customer Journeys” to align and link cross unit activities in their Operating Models — such User/Customer Journeys can be considered to be instance models of Contract execution, enabling Enterprises to assess and select appropriate features in their Operating Models.

standardization: Trade-offs between Individual Efficiency and Overall Effectiveness in the business processes A key challenge many enterprises face is in the choice between improving efficiencies focused on individual business processes and effectiveness of the overall end-to-end business flow. In the introductory section, we have introduced the idea of using Business Identities as intuitive anchor for an Enterprise’s engagement with its stakeholders. To address the trade-off challenge on individual vs collective efficiency in its business processes, an Enterprise’s Business Identity can help in (a) aligning the guidance of the enterprise’s vision/mission/values to the trade-off and (b) ensuring consistency in practices and tools, through standardization, in the execution of business processes.

In order to best support a company’s strategy, a foundation for execution is necessary. Such a foundation is the operating model of an organization.

See appendix Business Operating Model on page 85 for a summary of the theory around business operating models.

Contribution to the Enterprise

The operating model could help organizations define their requirements and expectations in terms of standardization and integration across business units. This would be useful to harmonize business capabilities and glossaries across companies and their business units.

Furthermore, an operating model may serve as input to the development of an Enterprise Architecture (see **??**) that serves as the organizing logic for business processes and IT infrastructure based on the company’s standardization and integration requirements.

Contribution to the EKG

Since an operating model reflects the business needs for standardization across a company and its business units, it could be a great input for the development of reusable business vocabularies as part of the data infrastructure of an company as later explained in the data part of this work. These vocabularies may, in turn, be converted to machine-readable models. Such machine-readable models are called ontologies.

A.3.2 Performance Management

Description

Measure business processes to manage performance.

> *"Enterprise Performance Management (EPM) — The process of monitoring performance across the enterprise with the goal of improving business performance. An Enterprise Performance Management (EPM) system integrates and analyzes data from many sources, including e-commerce systems, front- and back-office applications, data warehouses, and external data sources. Advanced EPM systems can support many performance methodologies, such as the balanced scorecard. "*

> *"Making a distinction between strategy and execution can do great damage to a corporation."*
>
> Martin [Mar10]

Performance Management at enterprises, for various reasons of convenience and expediency, has operated at (at-least!) three different levels:

1. **Periodic reporting cycles used by the executive leadership in evaluating progress on strategic goals**
   - These have been traditionally dominated by lag indicators such as receivables, payables, market share with intuition and judgment driven evaluation of lead indicators such as outages, downtimes, revenue pipeline, attrition etc. The Business Intelligence movement and the rapid digitalisation of key activities has improved the ability of most enterprises to collate, aggregate and report on various metrics of interest in such review cycles.
2. **Operations and Delivery management reporting used by front-line and operations staff to identify, analyse and fix problems and deviations**
   - Improvements in the planning and scheduling activities, through multiple generations of resource planning and workflow management, have helped most enterprises to rapidly configure chains of activities, optimise resource capital and monitor in near real-time the fulfilment and deviations in operations.
3. **Personnel evaluation and reporting towards role management for career and aspirations management of staff**
   - Performance appraisal has received a lot of attention in recent years, with shifts from intuition and heuristics based once-a-year activity to statistical profiling mechanisms and of late, continuous conversations focused on metrics such as productivity, retention and capability development.

While the Operating Model can address many of the disconnects that arise between these levels, enterprises still struggle with bridging the gap in aligning performance management for reasons such as:

- **Differences in focus between governance and operating procedures:**
  - Metrics in governance focus on enforcement whereas metrics in operations focus on output and progress. The feedback mechanisms which aim to align the two are dominated by subjective assessments resulting in high exception workloads.
- **Fulfilment focused integration leaving significant silos between information technology and operations**
  - Most use of information technology is focused on information flows and aggregation whereas operations technology is focused on individual units of activity. The relative prioritization is dominated by committee-based subjective assessments resulting in high backlogs, long wish lists of technology enablement and proliferation of tactical approaches in operations.
- **Unclear alignment of incentives, especially between operations and role performance**
  - While performance appraisals and role management have become increasingly automated in evidence management with links to operations and competency development, work assignment and activities seldom remain aligned to start-of-period goals, leading to subjective adjustments through mechanisms such as continuous conversations. The key challenge of mapping enterprise strategy to role specific goals remains mostly a subjective and judgmental exercise.

[1] https://ASCM.org

Contribution to the Enterprise

While practices around Performance Management might vary across enterprises, we highlight below a few focus areas where EKGs can act as effective enablers:

Business Identities can act as the common contexts for aligning strategic and operational performance management. Highly engaged senior executives practice these links to drive effective alignment with feedback through their own mental models of linkage. EKG driven feedback mechanisms, based on business identities, can then replicate such links and operate with more objectivity by enabling decision-making through links of operational activities and their metrics to the associated business goals, values and customer propositions.

Contracts can act as common contexts for aligning operations and role performance management. Employee contracts can be the most concrete form of expectation specification and management and their linkage to operations and the associated propositions of the enterprise to its customers and other stakeholders can help in aligning appraisals more objectively to operations and enterprise performance.

Contribution to the EKG

We welcome your input here: The Performance Management capability contributes to the EKG as follows...

A.3.3 Risk Management

Description

Identify risk to mitigate or accept, with reaction plans.

> *"Risk management — The identification, assessment, and prioritization of risks followed by coordinated and economical application of resources to minimize, monitor, and control the probability and/or impact of unfortunate events or to maximize the realization of opportunities. "*
>
> - ASCM[2]

Risk Management has traditionally been an integral, and sometimes implicit, part of an Enterprise's Business Governance practices. Though there are guidelines for business governance in every country, not all are formulated as regulations; some are only generally accepted norms of conduct[ENISA]. Some examples of well-known governance and risk frameworks are

- The BASEL Accord framework (for financial services organizations) used by regulators globally for their specific jurisdiction directives covering management practices around operational, credit and market risks
- COSO, a private sector initiative (sector agnostic) dedicated to helping organizations improve performance by developing thought leadership that enhances internal control, risk management, governance and fraud deterrence

Enterprises use a variety of internal control systems, both formally and informally, to monitor their governance mechanisms. The COSO[3] model[ERM-COSO], for example, defines internal control as *...designed to provide reasonable assurance of the achievement of objectives in the following categories:*

- *Operational Effectiveness and Efficiency*
- *Financial Reporting Reliability*
- *Applicable Laws and Regulations Compliance*

A recent study[2021-Risk] on the practices around Enterprise Risk Management (ERM) has highlighted the following:

- Effective risk management is a priority among boards of directors; organizations are facing pressures from a number of stakeholders to provide more risk information and business leaders want to be better prepared when unexpected risk events emerge to avoid being surprised
- Most organizations continue to struggle to integrate their risk management and strategic planning efforts; risk tolerances are not formally articulated as part of their strategic management activities

[2] https://ASCM.org
[3] Committee of Sponsoring Organizations of the Treadway Commission (COSO)

- There are a number of impediments to advancing an organization's risk management processes, with the belief that "risks are managed in other ways besides ERM" dominating the list
- There is a heavy emphasis on risks related to technology, legal/compliance and financial issues, with ERM processes less focused on emerging strategic/market/industry risks or risk related to reputation
- There are opportunities to reposition an entity's risk management process to ensure risk insights generated are focused on the most important strategic issues

Contribution to the Enterprise

Considering the above on the state, imperatives and opportunities in an enterprise's risk management practices, we highlight below a few key focus areas where EKG can help/enable enterprises:

Holistic, inter-related, strategy aligned and performance monitoring integrated risk management The topical approaches to risk management is a direct result of the silos of operations in various enterprise units. As outlined in the earlier sections, through Business Identities, EKGs can enable enterprises to align business strategy with the operating model and the performance management, thus creating bridges between and aligning the silos of operations to enterprise strategy. The inter-relationship, in the context of strategic objectives, can enable risk management to evaluate and monitor risk in a holistic manner. Qualitative and subjective metrics to more quantitative and objective metrics based risk management practices: Narrowly focused risk factors, owing to siloed specialisations have led to use of predominantly qualitative metrics in risk monitoring. EKGs can help align risk factors with metrics in performance management, thus helping in use of more effective quantitative objective metrics in risk monitoring.

Evaluation and prioritization of emerging risks through explorative analysis: Specialised areas of risk dominate focus of risk management, while strategic issues require a coordinated action across multiple operational areas. In helping to align risk monitoring more closely to the strategic issues, EKGs can enable enterprises to evaluate and prioritize key emerging risks affecting the strategy through enterprise wide, exploratory scenario-based analysis.

Contribution to the EKG

We welcome your input here: The Risk Management capability contributes to the EKG as follows...

A.3.4 Supply Chain Management

Description

Managed flow of goods, services & information between businesses.

> *"Supply chain management — The design, planning, execution, control, and monitoring of supply chain activities with the objective of creating net value, building a competitive infrastructure, leveraging worldwide logistics, synchronizing supply with demand, and measuring performance globally. "*
>
> - ASCM[4]

Contribution to the Enterprise

In the EKG context, the above Association for Supply Chain Management (ASCM) definition of *Supply Chain Management* could be broadened by also including any detailed supply chain within an ecosystem or enterprise where any one given "operation" delivers "value" to another "operation".

Contribution to the EKG

We welcome your input here: The Supply Chain Management capability contributes to the EKG as follows...

[4] https://ASCM.org

A.3.5 Capability Map

Description

Defining what a business does and which outcomes it wants to achieve.

In recent years, the word capability has steadily been drawing attention in the business world as a way to narrow the gap between business and IT. A capability is described by Ulrich and Rosen, (p.1, 2011) as:

> *"A business capability, or simply a "capability," defines what a business does. It does not communicate or expose where, why, or how something is done — only what is done. Specifically, the business capability is "a particular ability or capacity that a business may possess or exchange to achieve a specific purpose or outcome."*

Ulrich and Rosen [buscap]

Understanding *what* a business does is at least as important as understanding *how* it does it. Focusing on the *what* rather than on the *how* often provides the right level of abstraction of complex ecosystems in a way that can be easily digested by business executives and planning teams. Dissecting a business as a set of basic capabilities allows one to view complex organizations and business environments in a wide range of different ways. Additionally, one could zoom in on lower levels of abstraction to reach higher levels of detail depending on the needs of the target audience.

Consider the example of the high-level capability called *customer 360°*. This capability refers to having a complete picture of the customer and this is a capability many organizations may benefit from. A customer may be called a taxpayer in the context of a government institution, a patient in the context of a hospital and a student in the context of a university. Being able to manage different aspects of a customer might be imperative for many businesses depending on their business strategy. For instance, *customer 360°* as a high-level capability could entail the *customer trend analysis* and *customer management information* capabilities as lower-level capabilities.

In other words, this means that capabilities are *concerns* that could be refined into lower-level capabilities that in turn could be further refined. The advantage of functional decomposition in this hierarchical and modular way is that it increases reusability of capabilities across business lines, enterprises or even industries. Additionally, decomposing capabilities in different levels of abstraction offers a better sense of how c apabilities fit in on the overall view of the enterprise.

Capabilities should follow certain principles as follows:

1. **Capabilities define what a business does and not how a business does it.** This implies that a capability does not refer to a process or value stream.
2. **Capabilities are nouns, not verbs.** To make the distinction between a capability and a process, capabilities should be named with nouns such as customer management instead of a verb managing customers.
3. **Capabilities should be defined in business term and not in technical terms.** Executives and other business professionals should be able to clearly understand what a capability means — and take ownership of it — without being bothered with the technical details of the systems that implement it.
4. **Capabilities should be stable (i.e. least volatile) as possible.** There is a fundamental set of high-level capabilities that are necessary to run a business regardless of the sector of industry it is located in. Additionally, high-level capabilities may rarely change within an organization unless there are significant changes to its business strategy or organization.
5. **Capabilities are not redundant.** Capabilities appear once and only once on a capability maps even though they are shared by more than division, business line or business unit of an enterprise.
6. **There should only be one business capability map per business.** In order to bring transparency across the organization, there should be a single unified and harmonized business capability map per business. Therefore, business capabilities spread throughout different business units, divisions, department, lines of business, etc should be only appear once in the business capability map.
7. **Capabilities map to — but are not the same as — business processes, business units, LOBs and value streams.** There is rarely a one-to-one mapping between a business process, a business unit, an LOB or a value stream, and a capability. Therefore, mistaking a capability with one these other concepts may result in an overly complex overview of things that are constantly repeated across the enterprise. Additionally, we're trying to understand the what rather than the how and this would not be possible if capabilities are mistaken with process-related concepts.

8. **Automated capabilities are still business capabilities and not IT capabilities.** Capabilities are owned by the business as they are made by and for the business. Therefore, a capability that has been automated with an IT system is not an IT capability. Instead, it is nothing else than an *automated business capability*.
9. **Capabilities are the most valuable when incorporated in the overall pictures of a business' ecosystem.** While being useful in a stand-alone basis for discussion and planning, capabilities fulfil their biggest potential by being combined into a larger business capability map that represents the business ecosystem of an enterprise.

Contribution to the Enterprise

As a business/IT transformation enabler, capabilities may provide certain benefits to enterprises:

- **Capabilities provide business with a common language.** Capabilities offer business professionals and C-level executives a common understanding of which areas of an enterprise they need to address. The larger the enterprise, the more useful this idea is since business issues and strategy may revolve around many business units, lines of business, business processes, or even enterprise boundaries.
- **Capabilities may enable more precise investments.** Measuring the ROI of investment projects can be very difficult as the results might be spread across a myriad of different LOBs or business units. Instead of reconciling results from all these different places, one could merely look at the resources being allocated at a certain capability and this may allow business executives to have better estimates of the impact of their investments on the enterprise as a *whole*.
- **Capabilities serve as a baseline for strategic planning, change management and impact analysis.** By shifting the focus from business units, lines of business, IT systems and business processes to business capabilities, one is able to increase the traceability of strategic decisions into the daily operations and the business performance of the enterprise as a whole. So, capabilities serve as a starting point for tracking the impact horizontal and vertical impacts of strategic and tactical decisions from the executive team.
- **Capabilities lead to better business service specification and design.** Capabilities provide business-focused abstraction of the functionalities and information that information systems must provide. For instance, such capabilities may help in the selection or construction of services in a service-oriented architecture (SOA) as they directly embody business requirements.

In a few words, capabilities provide business executives the possibility to cut through the complexity inherent to most enterprises in order to make sounder decisions for strategic planning, impact analysis and investment planning. Using capabilities enables business to make decisions based on *what* needs to be resolved without initially dealing with the *how*. At the same time, it provides a way to tie the *how* to the results of the *what* for further validation that business efforts are delivery the expected results and for better alignment. On the other hand, the alternative to this approach involves repeating the traditional silo-focused approach of looking at hundreds or even thousands of IT systems, before being able to look at actual business actions and business results.

Contribution to the EKG

The *Business Capability Map* enables an enterprise to structure their Enterprise Knowledge Graph. Each capability is a candidate to become a *use case* for the EKG.

A well-designed mature EKG is a facade in front of all technical and organizational silos which means that — without any further structure — users do not experience these silos anymore. In many ways that may be a good thing but having silos — or different perspectives — can also be necessary. The EKG allows an organization to rethink their silos without being held back by their current data or technology landscape (or "technical debt").

The *Business Capability Map* is the ideal initial structure for the new EKG "silos". Business Capability Maps are usually rather coarse-grained and visualized in a three-level hierarchy whereas the structure of the EKG goes much further than that. Translating each capability to a "use case" is a good start but each of these use cases can be further broken down into smaller use cases where each use case becomes a highly reusable component of the EKG. This leads to a hierarchical structure, a taxonomy so you will, of all use cases, also called "the Use Case Tree (UCT)". This use case tree, at the higher levels, corresponds with the Business Capability Map and allows a business and its executives to "own" and control all the various parts of their EKG.

Part B

Organization Pillar

The Organization Pillar has the following components:

- B.1 Executive Leadership
- B.2 Product Ownership
- B.3 EKG Delivery Management
- B.4 Organizational Culture
- B.5 Organizational Capabilities

Chapter B.1

Executive Leadership

The Business Strategy Actuation component has the following capabilities:

- B.1.1 Organizational Strategy — Organizational mandate and strategic support for the EKG.
- B.1.2 Organizational Policies — EKG alignment with the views and values of the organization.
- B.1.3 Funding & Resources — Funding process and resource allocation for the EKG.
- B.1.4 Technology Requirements — CIO support for EKG (where EKG fits in the overall technology stack).
- B.1.5 Organizational Processes — Alignment of EKG with business architecture.
- B.1.6 Organizational Governance — Process for aligning EKG with organizational accountabilities.

B.1.1 Organizational Strategy

Description

Organizational mandate and strategic support for the EKG.

We welcome your input here.

B.1.2 Organizational Policies

Description

EKG alignment with the views and values of the organization.

We welcome your input here.

B.1.3 Funding & Resources

Description

Funding process and resource allocation for the EKG.

We welcome your input here.

B.1.4 Technology Requirements

Description

CIO support for EKG (where EKG fits in the overall technology stack).

We welcome your input here.

B.1.5 Organizational Processes

Description

Alignment of EKG with business architecture.

We welcome your input here.

B.1.6 Organizational Governance

Description

Process for aligning EKG with organizational accountabilities.

We welcome your input here.

Chapter B.2

Product Ownership

The Business Model Elaboration component has the following capabilities:

- B.2.1 Use Case Requirements — Identification/prioritization of EKG use case requirements.
- B.2.2 Funding & Budget — Annual and multi-year funding process.
- B.2.3 Delivery Process — Management of statements of work for business outcomes.
- B.2.4 Measurement Criteria — Business metrics and criteria for evaluating the EKG.
- B.2.5 Team Requirements Management — Identification and management of skill sets needed for the EKG.

B.2.1 Use Case Requirements

Description

Identification/prioritization of EKG use case requirements.

We welcome your input here.

B.2.2 Funding & Budget

Description

Annual and multi-year funding process.

We welcome your input here.

B.2.3 Delivery Process

Description

Management of statements of work for business outcomes.

We welcome your input here.

B.2.4 Measurement Criteria

Description

Business metrics and criteria for evaluating the EKG.

We welcome your input here.

B.2.5 Team Requirements Management

Description

Identification and management of skill sets needed for the EKG.

We welcome your input here.

Chapter B.3

EKG Delivery Management

The Business Enablers component has the following capabilities:

- B.3.1 ETL/Data Movement — Process for extracting/transforming data from source-to-EKG.
- B.3.2 Ontologies & Mapping — Approach to concept modeling and transformation.
- B.3.3 DataOps Process — Rules and standards for building, testing, and maintaining the EKG.
- B.3.4 User Interface & Access — Methodology for facilitating end-user access to data.
- B.3.5 EKG/Platform Team — Requirements and accountabilities for managing the EKG/Platform.

B.3.1 ETL/Data Movement

Description

Process for extracting/transforming data from source-to-EKG.

We welcome your input here.

B.3.2 Ontologies & Mapping

Description

Approach to concept modeling and transformation.

We welcome your input here.

B.3.3 DataOps Process

Description

Rules and standards for building, testing, and maintaining the EKG.

B.3.4 User Interface & Access

Description

Methodology for facilitating end-user access to data.

We welcome your input here.

B.3.5 EKG/Platform Team

Description

Requirements and accountabilities for managing the EKG/Platform.

We welcome your input here.

Chapter B.4

Organizational Culture

The **??** component has the following capabilities:

- B.4.1 Implementation Approach — Build versus buy decisions.
- B.4.2 Agility & Innovation — Empowerment, adaptability and failure tolerance.
- B.4.3 Ecosystem Collaboration — Alignment with technology, business, and operations.

B.4.1 Implementation Approach

Description

Build versus buy decisions.

We welcome your input here.

B.4.2 Agility & Innovation

Description

Empowerment, adaptability and failure tolerance.

We welcome your input here.

B.4.3 Ecosystem Collaboration

Description

Alignment with technology, business, and operations.

We welcome your input here.

Chapter B.5

Organizational Capabilities

The Organizational Capabilities component has the following capabilities:

- B.5.1 Control Functions — Coordination with privacy, legal, security, records management.
- B.5.2 Vendor Management — Acquisition and licensing of vendor products.
- B.5.3 Talent Management — Skills acquisition, retention and training.
- B.5.4 Knowledge Sharing — Ability of the organization to facilitate knowledge collaboration.

B.5.1 Control Functions

Description

Coordination with privacy, legal, security, records management.

We welcome your input here.

B.5.2 Vendor Management

Description

Acquisition and licensing of vendor products.

We welcome your input here.

B.5.3 Talent Management

Description

Skills acquisition, retention and training.

We welcome your input here.

B.5.4 Knowledge Sharing

Description

Ability of the organization to facilitate knowledge collaboration.

We welcome your input here.

Part C

Data Pillar

The Data Pillar has the following components:

- C.1 Data Strategy
- C.2 Data Architecture
- C.3 Data Quality
- C.4 Data Governance

Chapter C.1

Data Strategy

The Data Strategy component has the following capabilities:

C.1.1 Goals & Objectives — EKG as the semantic data mesh for the organization.

C.1.2 Business Case — Business rationale, justification and ROI of the EKG.

C.1.1 Goals & Objectives

Description

EKG as the semantic data mesh for the organization.

We welcome your input here.

C.1.2 Business Case

Description

Business rationale, justification and ROI of the EKG.

Chapter C.2

Data Architecture

The Data Architecture component has the following capabilities:

C.2.1 Ontologies — The *models* that provide the structure and meaning for the information.

C.2.2 Business Vocabularies — Terms used to refer to the information by different business communities.

C.2.3 Datasets — Coherent collections of information made available for access.

C.2.4 Data Mapping — Unifying or relating different information concepts.

C.2.1 Ontologies

Description

The models that provide the structure and meaning for the information.

An ontology is a set of formal definitions for the key concepts that organize and structure an organization's information. Having ontologies provides a *common denominator* level of unification that allows information to be shared and communicated between different use cases and shareholders, regardless of the different sources, structures and vocabulary they might use.

An ontology is about:

- explicit meanings and relationships; the terms used are less important
- a combination of definitions in both text and logic

An ontology can be the basis of, but is broader than:

- a taxonomy
- a vocabulary
- a data or object model
- a conceptual model
- a specific serialization format

Ontologies can be expressed at different levels of sophistication, with different scopes, and in a combination of languages. The basic structures include:

individual a representation of a business object or item which is the subject of information to be managed. An individual usually has a unique identity. For example `Person X` or `Shipment Y`. Many such individuals might represent the same real world object.

data value strings, numbers, dates which represent the data.

property a type of data point that may be associated with individuals. An individual, a property and a value — which may be a data value or another individual — form a triple.

For example person `X hasBirthDate D` or person `X hasMother Y`. Triples whose value is another individual form relationships. Properties may have generalizations, for example `hasMother` is a `subProperty` of `hasParent`.

class a category applied to individuals, that determines what you can do with them, the properties you can expect to see, and the rules that might apply; an individual may be a member of many classes associated; classes may have generalizations. Note that, unlike more traditional approaches, properties are independent of classes though they may be used to infer class membership. For example, given the triple `X hasMother Y` you may be able to infer that both `X` and `Y` are members of the class `Person`, or at least `Animal`.

ontology grouping of the above for management and identification purposes.

Contribution to the Enterprise

Ontologies are needed to truly understand what a given set of data really means and what can be inferred from it. For example you cannot rely on the name of a column in a spreadsheet. A deceptively simple column name such as "number of European customers" leaves open the meaning of "European" and "customer" and timing (when does one start and stop being a customer?). And different sources could have different interpretations of that same name. The benefit is consistency, accuracy and the ability to make sound business decisions. Having the models themselves be resources that can be looked up means that all data is self-defining and carries its meaning with it. In an EKG there is no fixed set of ontologies so it can non-disruptively incorporate additional knowledge. Ontologies allow data to be understood independent of the format/technology and the vocabulary used in different communities, saving misunderstandings and battles over which word to use.

Contribution to the EKG

Ontologies are the basis for Principle 2 — 'Meaning' on page 77:

> *"The meaning of every data point must be directly resolvable to a machine-readable definition in verifiable formal logic."*
>
> - EKGF[1]

The link to precise meaning serves to mitigate problems created using the same word with multiple definitions; and the challenges of expressing conceptual nuance using multiple informal sentences. In the other direction, from ontology to vocabulary, it should be possible to generate a business glossary directly from ontologies for a given scope. Since they should capture the meaning of concepts applicable to an organization, or an even broader ecosystem, the choice of concepts to include in an EKG should be driven by business use cases. And different overlapping ontologies may be included and mapped to cover different relevant aspects. Likewise, it should be possible to generate — and map to — models for more conventional tools from ontologies, by applying technology-specific rules.

Semantic modeling also eliminates the problem of hard-coding assumptions about the world into a single data model. And while multiple ontologies may coexist, they are able to be mapped and connected to each other. In a mature environment, the data modeling process drives technology implementation, by defining the detailed data structures and associated Application Programming Interfaces (APIs). These components — along with functional code — are included as part of the testing suite within the EKG to facilitate rapid deployment.

Different types of external data models are not needed in EKG but can be mapped to or generated. In fact, physical data modelers are a community with their own vocabulary.

Constraints/shapes for models are applied by context (use case) — there is no Single Version of (the) Truth (SVOT) for the EKG as a whole. Different ontologies may be used for different contexts and mapped to each other in the underlying knowledge graph.

Ontology elements are linked to by vocabularies (see C.2.2) and mapped to other data models and datasets (C.2.3) to provide their meaning; and from Use Cases to provide their scope. These aspects are covered by those respective capabilities.

Dimensions

C.2.1.1 How much of the enterprise data is covered by ontologies?

[1] `https://www.ekgf.org/principles`

C.2.1.2 How well is ontology coverage mapped to business need?

C.2.1.3 To what extent are concepts independent of but mapped to terminology/vocabulary?

C.2.1.4 Level of sophistication of textual and logic definitions

C.2.1.5 Level of tooling is available and used

C.2.1.6 Level of training and trained people

C.2.1.7 Level of process (including change management), guidelines and standards

C.2.1.8 Level of modularity and reuse — internal and external

C.2.1.9 Extent of examples and tests

C.2.1.10 Extent of traceability with different logical and physical data models

Levels

The following criteria for each level are abbreviated: each item is shorthand for:

- documented process
- trained participants
- implemented process and/or technology
- monitoring and improvement

Levels 1–5

Level 1: EKG Initiation

- ☐ Minimal ontologies which could be as simple as a list of classes and properties used in graphs
- ☐ Basic metadata (definition, label) for each class and property
- ☐ Each individual (in data) has at least one explicit class
- ☐ Ontology coverage for each use case in scope of the project; project includes minimal number of ontologies and classes not justified by a use case
- ☐ Definitions catalogued and under change management

Level 2: Extensible Platform

- ☐ Ontologies expressed in a standard ontology language (could be as simple as RDF Schema)
- ☐ Common (shared or mapped) concepts across EKG projects
- ☐ Ability to see ontology usage by use cases, vocabularies and datasets
- ☐ Namespace scheme established and used for new ontologies in the EKG
- ☐ Ontology guidelines in place and implemented, including common metadata
- ☐ Documented approach for external ontologies, including selection and adaptation
- ☐ Annotated example files for documentation and training
- ☐ Test files based on use cases covering all used ontology elements
- ☐ Ontology change management includes impact analysis and stakeholder approval
- ☐ Tooling for ontology diagrams and documentation
- ☐ Automated basic checking of ontology syntax
- ☐ Access to at least one trained ontologist

Level 3: Enterprise Ready

- ☐ Modeling of required data and constraints by use case, including for stored and communicated data
- ☐ Automated validation of ontologies (for guideline compliance, and for logical consistency), with results as triples
- ☐ Automated testing and validation of test data with ontologies (per use case)
- ☐ Separation of concerns to support enterprise management such as bi-temporality, transactions and events
- ☐ Automated transformation of ontologies to use common serialization and metadata
- ☐ Automated checking of ontologies against different profiles (e.g. OWL-RL) to check for technology support
- ☐ Automated checking of ontologies against different best practices
- ☐ Ontology source changes linked to automated EKG Operations (see D.2.2) for testing and deployment
- ☐ Impact analysis identifies ontology breaking changes which require fixes to existing EKG data
- ☐ EKG-wide ontology browsing and searching
- ☐ *Follow-your-nose* UI starting from any ontology element [2]
- ☐ *Follow-your-nose* API starting from any ontology element [2]
- ☐ Trained ontologist available to each project (ideally via the CoE for the EKG)

Level 4: Strategic Asset

- ☐ Separation of ontologies from vocabularies, with multiple vocabularies for different communities mapped to the same concepts
- ☐ Ontology architecture management process, including use of patterns and modularity
- ☐ Generation of logic into business language
- ☐ Automated fixes to existing EKG data in response to ontology breaking changes
- ☐ Basic ontology metrics and reporting, including usage in data
- ☐ Generation of ontologies/shapes for external interchange

Level 5: Operational Ecosystem

- ☐ Sophisticated ontology metrics and reporting, including trends
- ☐ Matching and differencing of ontologies from different sources
- ☐ Automated matching of ontologies with vocabularies
- ☐ Generation of validation code for external interchange
- ☐ Wizard for developing ontologies from business questions
- ☐ Inducing of ontologies from instance data

C.2.2 Business Vocabularies

Description

Terms used to refer to the information by different business communities.

Business terminology provides non-technical descriptions of an organization's processes, use cases and data assets. Mapping to ontologies provides the meaning and harmonizes the terms, providing transparency into definitions that helps promote consistency of business terms, identify synonyms and link business content to technical definitions. Agreement on granular meaning by involved stakeholders is essential to managing conceptual nuance across shared applications and linked processes. The business terms used are determined by the EKG use case or context.

[2] see also predicate-IRI

Contribution to the Enterprise

Parts of large organizations, and the others they communicate with, often have their own local business language. Likewise, operational systems — including their data warehouses and content repositories — use data models and predefined physical schemas that are designed to serve specific purposes. The ability of the knowledge graph to link this terminology with its actual meaning defined via ontologies at its most granular (atomic) level eliminates confusion over meaning when data is validated or transformed. Standalone business glossaries provide useful input but the terms need to be reconciled to their enterprise-level meaning via the EKG.

Contribution to the EKG

We welcome your input here: The Business Vocabularies capability contributes to the EKG as follows...

Dimensions

C.2.2.1 The organization has one (or many) officially designated glossaries

C.2.2.2 Level of review by Subject Matter Experts (SMEs)

C.2.2.3 Level of change management including versioning, approval, audit trail

C.2.2.4 Level of communication and ready availability

C.2.2.5 Terms have a full business definition

C.2.2.6 Business definitions also use a vocabulary with cross references

C.2.2.7 Definitions traced to their source with change triggering

C.2.2.8 Level of examples including edge case instances, which can expose vagueness or ambiguity

C.2.2.9 All terms mapped to an ontology definition

C.2.2.10 Level of review by ontology experts

C.2.2.11 Level of review by application/data experts

C.2.2.12 Level of coverage of use cases

C.2.2.13 Relationship to official business glossaries or data dictionaries

C.2.2.14 Mapping to business models including processes, objectives, Key Performance Indicators (KPIs)

C.2.2.15 Segregation of vocabularies by community

C.2.2.16 Reuse of glossaries across communities

C.2.2.17 Support for different natural languages

C.2.2.18 Searchability

C.2.2.19 Completeness of reverse mapping from ontologies to in-scope glossary for terms and definitions

C.2.2.20 Generation of glossaries from ontologies

C.2.2.21 Relationship amongst terms (synonyms, abbreviations)

C.2.2.22 Export to alternative representations e.g. web page, spreadsheet

C.2.2.23 Managing proper nouns (names of business entities, products etc) and mapping to individuals in the EKG

Levels 1–5

Level 1: EKG Initiation

- [ ] Business terms for EKG use cases are captured and mapped to an ontology (possibly as simple labels)

Level 2: Extensible Platform

- [ ] Relevant terms (for EKG use cases) are associated with the use case independently of their corresponding ontologies

Level 3: Enterprise Ready

- [ ] Existing glossaries within the organization, within the scope of supported use cases, are mapped to ontologies and imported into the EKG
- [ ] Terms are grouped into vocabularies for reuse in different communities
- [ ] Vocabularies include local textual definitions as well as being mapped to ontologies
- [ ] Community level vocabularies able to import common vocabularies reusable in many communities: allow for multiple levels from local to global
- [ ] Support for synonyms and abbreviations within vocabularies
- [ ] Vocabularies use general principles of modularity
- [ ] Ontologies presented to different communities in their own language
- [ ] Natural language used for definitions links to other terms used
- [ ] Terms are searchable via the knowledge graph interface
- [ ] Generate vocabularies from existing data models or ontologies
- [ ] Ontology logic is presented to business as natural language

Level 4: Strategic Asset

- [ ] Ontology logic is presented to business as natural language
- [ ] EKG is the authoritative source for all terms, scoped by community, context and use case
- [ ] The governance processes for all new terms (and changes) are managed directly within the EKG
- [ ] The results of ontology inferencing are presented in business natural language
- [ ] Use of ontology logic to validate use of vocabulary terms
- [ ] Management of homonyms and disambiguation of terms — through being mapped to different concepts
- [ ] Use of business terms in natural language queries
- [ ] Techniques for disambiguation e.g. allowing business users to select the intended meaning
- [ ] Use of textual analytics to suggest modifications to vocabulary/ontology (lemmatization)
- [ ] Use of vocabularies to match internal databases or external data sources for linked data, and establish `sameAs` links

Level 5: Operational Ecosystem

- [ ] Terminology is used to support Natural Language Processing (NLP) of unstructured data for the EKG
- [ ] Use of reference data from KG to support entity resolution

C.2.3 Datasets

Description

Coherent collections of information made available for access.

A Data Inventory is a repository of Data Assets for the organization (also known as a Data Catalog or Metadata Repository). The inventory contains information about what data exists, where it resides ("data-at-rest"), responsible parties, upstream/downstream usage, how it moves ("data-in-motion"), classifications, quality designations, availability, and other useful metrics. The contents are cataloged at physical, logical, and business levels using defined organizational standards. Consistently documenting data is the backbone of an effective data management program. Policies and controls are required to ensure the inventory remains accurate and relevant.

Contribution to the Enterprise

Linking data inventory to the knowledge graph (and business concepts) ensures the precision of meaning at the most granular level. Data in the knowledge graph is traceable to all application usage allowing users to find data of interest through assisted and contextual search and navigation. The unique ability of the knowledge graph to connect data with metadata enables users to perform flexible queries in ways that were not previously possible.

Contribution to the EKG

We welcome your input here: The Datasets capability contributes to the EKG as follows...

Dimensions

- C.2.3.1 Do one or more data inventories exist
- C.2.3.2 Is the inventory based on defined standards (for both meaning and format)
- C.2.3.3 Is defined and in-scope data (both breadth and depth) covered in the inventories
- C.2.3.4 Is the data inventory linked to Systems of Record (SORs) and authorized data distribution points
- C.2.3.5 Is the inventory linked to the business meaning of the data and expressed using standards
- C.2.3.6 Is the creation and maintenance of the data inventory mandated by policy and incorporated into the data strategy
- C.2.3.7 Is the quality of the content in the inventory measured, reported to involved stakeholders, and used for process enhancement

Levels 1–5

(concepts are inherited as levels progress)

Level 1: EKG Initiation

- ☐ Sources, data sets, and metadata are onboarded and expressed as formal ontologies
- ☐ Authoritative (upstream) data sources and (downstream) consumers are documented and verified by users, data, and technology
- ☐ The inventory of applications is defined and selected for graph applications
- ☐ Requirements and dependencies for each outbound data flow are documented and verified (implementation in the graph is not a requirement)
- ☐ Business glossaries for in-scope use cases are defined and verified in the graph (including a list of data sources and datasets)
- ☐ Policy implemented mandating inventory maintenance and only authorizing the use of data that has been logged into the inventory

Level 2: Extensible Platform

- ☐ UCTs are defined, standardized, and implemented
- ☐ All upstream data sources are linked to the authorized systems of record and distribution points
- ☐ Policy mandating the use of SORs and documentation of data flow is implemented
- ☐ Entitlements have been defined in the graph (governing access to sources of data in the inventory)
- ☐ Classifications (i.e. criticality, security, privacy) are aligned with the use case tree and captured in the knowledge graph
- ☐ Governance requirements (i.e. use cases, accountability, data sources, data flows, Service-level Agreements (SLAs)) are modeled and registered into the knowledge graph

Level 3: Enterprise Ready

- ☐ Data inventory is centralized in the graph and linked to governance for defined use cases
- ☐ Ontologies and data models (including change history and transformations) are registered in the knowledge graph
- ☐ Entitlements are calculated within the inventory and enforced at the datapoint level
- ☐ Data Quality is automatically calculated (fine-grained with dynamic value resolution) within the inventory for each use case
- ☐ Data retention rules are registered in the graph and automatically enforced
- ☐ Full audit trail for all upstream and downstream data usage is registered in the graph
- ☐ Data elements, calculation methods, and Critical Data Elements (CDEs) are linked to individual regulatory requirements

Level 4: Strategic Asset

- ☐ Connected inventory has been extended to include real-time (transactional) data
- ☐ Inventory is extended to external suppliers and third parties along the supply chain
- ☐ The inventory is fully integrated with machine learning to optimize data flow
- ☐ The "value of data" is calculated and classified within the organizational inventory

C.2.4 Data Mapping

Description

Unifying or relating different information concepts.

Contribution to the Enterprise

Data integration is challenging because data from multiple sources can have different file formats, data structures, definitions, and contextual meanings. Using the EKG during data integration standardizes the meaning of data and makes the content understandable to both humans and machines. Embedding referenceable meaning into the data using machine-readable standards facilitates automatic validation and assurance of data quality. Registering data integration activities into the EKG generates full data transparency across linked processes. Finally, ontology-based metadata representations make it possible to embed business rules and accommodate different values, identities, and definitions that existed at various times in the entity lifecycle.

Contribution to the EKG

We welcome your input here: The Data Mapping capability contributes to the EKG as follows...

Dimensions

C.2.4.1 Are the data integration activities, their systems, repositories, and connections known and tracked

C.2.4.2 Are data integration activities linked to data inventory, business glossaries, and data models

C.2.4.3 Are all data integration input and output datasets documented, tracked, and governed

C.2.4.4 Are there reusable standards and defined business rules for performing data integration

C.2.4.5 Are data integration patterns, tools, and technologies defined, governed, and used

C.2.4.6 Has the firm established a central data integration function (i.e. integration Center of Excellence) to manage ETL across both internal and external data pipelines

The goal is not always a single source of data- but rather the ability to choose the right authoritative source for the appropriate context.

Levels 1–5

Level 1: EKG Initiation

- ☐ All data sources are identified and documented for in-scope use cases
- ☐ Do we know the authoritative source for each data set (should not be able to do integration without using approved authoritative source)
- ☐ Does everyone agree that we are using the right sources (the right source for every context) – link to governance
- ☐ Do we have an approved list of what each source feeds (precise description at the entity level that we can get from an approved source — must know if this is the primary source of the data per the use case context). For any given entity do I have all the potential sources and for a specific context do I know which is authorized.
- ☐ There is a defined governance process for change management and testing (clear picture of all the dependencies for data integration). If there are changes to authoritative sources — do we know the downstream implications (tracked and tested)
- ☐ Are all technology stacks known and supported by current teams (are all key systems under the management and governance of the organization — should not have ghost systems that are not controlled as part of the integration process)
- ☐ Entitlement policies and classification rules (i.e. security, PII, business sensitive) are defined and verified
- ☐ Data Quality requirements are defined, documented, and verified

Level 2: Extensible Platform

- ☐ All information (above) are identified, precisely defined, and on-boarded into the knowledge graph
- ☐ Able to do datapoint lineage (detailed and complete view of the data integration landscape)
- ☐ Start making the EKG the central point for data integration (the EKG becomes the Rosetta stone of integration) — onboard systems, convert to RDF, integrate into EKG (defined as the data integration strategy — not necessarily complete)
- ☐ All data sets that are on-boarded into the EKG are coming from the authoritative sources. There are no man-in-the-middle systems. The goal is direct from the authoritative source to the target system for in-scope use cases. Must get the most granular data directly from the authoritative sources.
- ☐ All datasets are ”self-describing datasets (SDDs)”.
- ☐ **Policy** — All data is obtained from the EKG as the authoritative source. Do not go directly to the originating source of the data.
- ☐ Entitlement policies and classification requirements are on-boarded into the EKG
- ☐ Data quality business rules are on-boarded into the EKG

Level 3: Enterprise Ready

- ☐ Data is precisely defined (granular level)- expressed as formal ontologies- and on-boarded into the EKG
- ☐ All data flows are modeled, defined, and registered in the EKG (full lineage in the EKG for all in-scope applications)
- ☐ Start to make the EKG the authoritative source (set-up to facilitate decommissioning of systems). The EKG is structured to become the “new” system for in-scope applications (as soon as all connections emanate from the EKG).
- ☐ Entitlements are automatically managed enforced

Level 4: Strategic Asset

- ☐ **policy** — All downstream client systems are using authoritative sources as the only source of information for in-scope datasets (EKG is in the middle of all data flows)
- ☐ All “cottage industry systems” are replaced by the EKG (and EKG is able to perform all the requirements of any system it replaced – reporting, entitlement, quality control)

Chapter C.3

Data Quality

The Data Quality component has the following capabilities:

C.3.1 Data Quality Business Rules — Strategy and approach for managing data quality.

C.3.1 Data Quality Business Rules

Description

Strategy and approach for managing data quality.

Data quality is a measurement of the degree to which any dataset is fit for its intended purpose. It is based on an understanding of application requirements and derived by reverse engineering of the data!production process. A data quality framework is an agreed methodology including operational controls, governance processes and measurement mechanisms. The framework is designed to support organizational priorities for data quality based on criticality and business value.

Contribution to the Enterprise

Data quality is an integrated feature of the EKG. Data quality rules are linked to structured business vocabularies to ensure that meaning is shared and not obscured by vague terms or cryptic codes. The logic of business rules and policies are captured and expressed as executable models and consistently enforced across all systems and processes. These quality constraints (models) allow firms to measure the quality of the data and perform verification across disparate systems. In the mature EKG, violations of logic or integrity are identified and prevented before data enters the system.

Contribution to the EKG

We welcome your input here: The Data Quality Business Rules capability contributes to the EKG as follows...

Dimensions

C.3.1.1 Are data quality business rules specified, formalized, and expressed in a standardized manner

C.3.1.2 Is there a clear line of business accountability (owned, funded, and governed) for the data quality rules

C.3.1.3 Is there a centrally managed repository of business rules

C.3.1.4 Is there a clearly defined mechanism for logging additions and performing updates

C.3.1.5 Are the business rules aligned with business applications and traceable to source systems

C.3.1.6 Are the data quality business rules automated and expressed in a machine-executable format

Levels 1–5

Level 1: EKG Initiation

- ☐ Data quality business rules (conditions) have been defined, documented and, verified by SMEs (process for evaluation and acceptance defined)
- ☐ Business rules are aligned with in-scope use cases and specific user stories
- ☐ Business rules are standardized and registered into a repository with a defined mechanism for logging additions and performing updates

Level 2: Extensible Platform

- ☐ A defined architecture exists to translate business rules into machine-executable code (some rules will be OWL expressions, some will be SHACL shapes and constraints, some will be translated into workflow logic)
- ☐ Business provenance and lineage are traceable across the data!supply chain and evaluated against defined business rules (all business rules must be traceable and understandable in context — must understand the purpose and importance of the rule)
- ☐ Business rules for in-scope use cases are implemented in the EKG

Level 3: Enterprise Ready

- ☐ **Metrics** — The measurement criteria are defined for data quality business rules (which rules are executed, how often, improvement)
- ☐ **Performance** — The value of business rules are related to business concepts (products, financial performance, organizational objectives) — able to trace the core relationship between the business objectives and the data quality business rules (correlation between rules and outcomes are known, able to be queried and traceable within the EKG)

Level 4: Strategic Asset

- ☐ Business rules are combined with AI capability for compliance (dynamic optimization of business rules)
- ☐ Model-driven (senior management can begin to optimize business objectives using business rules in the EKG — i.e. alignment of business rules with “what if” scenarios)

Level 5: Operational Ecosystem

- ☐ All business rules are driven by business objectives (objectives are in the EKG with appropriate scorecards)

Chapter C.4

Data Governance

The Data Governance component has the following capabilities:

C.4.1 Data-management Operating model — Structure and approach for the EKG Center of Excellence.

C.4.2 Data-management Policy — Policies, procedures, and standards for managing the data lifecycle.

C.4.3 Classification Management — Management of the data production and manufacturing process.

C.4.4 Risk & Control Environment — EKG access control mechanisms.

C.4.1 Data-management Operating model

Description

Structure and approach for the EKG Center of Excellence.

The people, processes, capabilities, and tools that define the role of data management in delivering value to an organization's customers.[1]

The operating model can help stakeholders understand the complexity of the data manufacturing process and how components relate to each other. It specifies the roles and responsibilities of the stakeholders involved in the data management program. It provides a framework and policies to help align governance concepts with operating requirements and organizational culture. An effective operating model can both describe the way the organization does business today ("as is") and communicate a vision of how an operation will work in the future ("to be").

Contribution to the Enterprise

The operating model for the CoE for the EKG establishes a new way of running the organization to enable companies to accelerate development and operate more efficiently. The EKG governance!framework focuses on combining new data capabilities (i.e. resolvable identity, standard ontologies, open standards) as part of an integrated process. A well defined operating model specifies architecture components that support flexible and reusable data to help organizations achieve improvements in revenue, customer experience and cost.

Contribution to the EKG

We welcome your input here: The Data-management Operating model capability contributes to the EKG as follows...

Dimensions

C.4.1.1 Have the underlying principles (and challenges) of data management been established and accepted by involved stakeholders

C.4.1.2 Does the Office of Data Management (ODM) have authority to enforce adherence to policy

C.4.1.3 Has the data management funding model been established and implemented

[1] There is also a capability called "Operating Model" in the Business Pillar that has a different meaning, see section *A.3.1, "Operating Model"*.

C.4.1.4 Has the organization identified and recruited stakeholders with sufficient skill sets to implement the data management program

C.4.1.5 Are involved stakeholders held accountable to data management program deliverables

C.4.1.6 Has the organization defined and validated the operating models, and workflows necessary to implement the data management program

C.4.1.7 Are metrics and KPIs captured and actively used to improve the data management operating process

C.4.1.8 Is the data management operating model audited for compliance and effectiveness

C.4.2 Data-management Policy

Description

Policies, procedures, and standards for managing the data lifecycle.

Data management policies are the mandated requirements to ensure effective management of an organization's information assets. Policies are implemented via operating models, standards and documented procedures. Data management policies define requirements related to data access (e.g. authorized sources, entitlements, usage and redistribution), data quality (e.g. metadata, definitions, modeling and business rules), organizational governance (e.g. accountabilities and alignment with business processes) and control functions (e.g. retention, security, privacy and change management).

Contribution to the Enterprise

Data management policies and standards should be crafted to cover the specifics of the knowledge graph. Candidate areas for EKG-related policies include those associated with the management of data catalogs, the resolution of identity and data meaning, the standards for data modeling and metadata, the use of SORs/authorized distribution points and the control over data entitlements.

Contribution to the EKG

We welcome your input here: The Data-management Policy capability contributes to the EKG as follows...

Dimensions

C.4.2.1 Are policies and standards in alignment with data management strategy

C.4.2.2 Are data management policies documented, complete and operational

C.4.2.3 Are data management policies linked to operational control functions as well as the SDLC process of the organization

C.4.2.4 Have data policies been created in collaboration with (and approved by) business, technology and operational stakeholders

C.4.2.5 Are data management policies aligned with the requirements of the EKG

C.4.2.6 Have data management policies been approved by executive management and audit

C.4.3 Classification Management

Description

Management of the data production and manufacturing process.

Data originates from a wide variety of business processes and flows through multiple systems before it is ingested into applications and analytical functions. As data flows across this "chain of supply" — it can be renamed, enriched, transformed and updated (many times). Data producers and consumers must work collaboratively to ensure that the data is fit for its intended purposes. This necessitates communications about the requirements for quality as well as an understanding of how the data is being used, where it originates, how it has been transformed and its consistency with original intent.

Contribution to the Enterprise

In the EKG, CDEs are defined by data quality business rules and expressed as machine-executable models. , These rules are automatically executed across systems, processes and applications to ensure consistency. Data is linked to the ontology and resolved to a single identifier to mitigate confusion about meaning when the data is onboarded or transformed. The EKG is able to trace data flow and verify that criticality is aligned with SORs and agreed usage.

Contribution to the EKG

We welcome your input here: The Classification Management capability contributes to the EKG as follows...

Dimensions

- C.4.3.1 Are critical data elements identified, verified and prioritized for specific use cases and applications
- C.4.3.2 How is the organization managing the distinctions between granular data attributes and derived/calculated business concepts
- C.4.3.3 Is the inventory of critical data elements implemented and linked to how the data is being consumed
- C.4.3.4 Are critical business elements connected to business glossaries, ontologies, SORs and authorized distribution points
- C.4.3.5 Is the front-to-back flow of data defined, validated and linked to the designations of critical data
- C.4.3.6 Are the governance mechanisms for managing critical data defined and operational

C.4.4 Risk & Control Environment

Description

EKG access control mechanisms.

Entitlement management is technology that grants and enforces access rights to data, devices, systems and services. Entitlement systems are linked to organizational policies (rules) governing access. These systems track access to applications to ensure that actions are in line with policies and to provide data about access for audit purposes. Entitlements need to be managed in sync with the goals for security, business continuity and compliance. There can be multiple proprietary systems that all have unique entitlement structures and individuals can move across departments and perform a variety of roles.

Contribution to the Enterprise

Control processes in the EKG environment are managed as a structured set of executable business rules or models — in the EKG, and automatically enforced across all use cases. The meaning of the data is structurally validated at the point of data capture to prevent bad data from entering the system at ingestion. In the knowledge graph, data and metadata are fully integrated to ensure reusability across systems and operational processes. The EKG is able to map data flow including all dependencies and transformations to verify that data is obtained from authorized SORs, identified according to access rights and aligned with intended usage.

Contribution to the EKG

We welcome your input here: The Risk & Control Environment capability contributes to the EKG as follows...

Dimensions

- C.4.4.1 Has the organization developed a structured framework outlining the principles of how operational risk is identified, assessed, monitored and controlled
- C.4.4.2 Has the operational risk framework been adopted by executive management and verified by internal audit
- C.4.4.3 Have oversight mechanisms been adopted to ensure compliance with operational risk control policies and governance procedures

C.4.4.4 Are technology resiliency and continuity plans in place to ensue systems integrity, security and availability during mergers, acquisitions and consolidations

Part D

Technology Pillar

The Technology Pillar has the following components:

- D.1 Technology Strategy
- D.2 Technology Execution
- D.3 User Interface

Chapter D.1

Technology Strategy

The Technology Strategy component has the following capabilities:

D.1.1 Technology Architecture — Plans for physical infrastructure, applications, and automation.

D.1.1 Technology Architecture

Description

Plans for physical infrastructure, applications, and automation.

Chapter D.2

Technology Execution

The Technology Execution component has the following capabilities:

- D.2.1 Data Integration — Ability to source and integrate data (consolidate silos).
- D.2.2 EKG Operations — Deployment and maintenance requirements for the EKG.
- D.2.3 Logging / Systems monitoring — Physical infrastructure and composite view of systems/applications.
- D.2.4 Business Continuity — Continued business operations during unplanned EKG/Platform disruption
- D.2.5 Data Quality Implementation — Technical implementation of data quality framework
- D.2.6 Analytics — Technical implementation of analytics facilities
- D.2.7 Inferencing Rules — Inferring new facts from existing facts by means of ontologies
- D.2.8 Entitlement Enforcement — Enforcing fine-grained context-specific entitlement policies
- D.2.9 Data Transformation Execution — Governed automated orchestration of data transformation
- D.2.10 Identity Resolution — Standards for ensuring content is identified and resolvable.
- D.2.11 Knowledge Graph Storage and Retrieval — Semantic data storage managed by the EKG/Platform
- D.2.12 Knowledge Graph Federation — Model-driven federated interoperable and distributed data platforms
- D.2.13 Knowledge Graph Virtualization — Model-driven views on data of any source leaving it in-place

D.2.1 Data Integration

Description

Ability to source and integrate data (consolidate silos).

D.2.2 EKG Operations

Description

Deployment and maintenance requirements for the EKG.

D.2.3 Logging / Systems monitoring

Description

Physical infrastructure and composite view of systems/applications.

D.2.4 Business Continuity

Description

Continued business operations during unplanned EKG/Platform disruption

D.2.5 Data Quality Implementation

Description

Technical implementation of data quality framework

D.2.6 Analytics

Description

Technical implementation of analytics facilities

D.2.7 Inferencing Rules

Description

Inferring new facts from existing facts by means of ontologies

D.2.8 Entitlement Enforcement

Description

Enforcing fine-grained context-specific entitlement policies

D.2.9 Data Transformation Execution

Description

Governed automated orchestration of data transformation

D.2.10 Identity Resolution

Description

Standards for ensuring content is identified and resolvable.

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

Scenario / Use Case- Use a URI for identification without considering impacts of impacts to unique identity by using an named identifier.

If you will only have a single gremlin vehicle, ever for-sale then you could go with a URI for which there isn't a concern for the uniqueness as it's not possible another will come along. Or, it could be managed manually by checking for existence of the name of the vehicle prior and incrementing.
First gremlin-
`<https://ekgf.org/cars/for-sale#gremlin>`
Second gremlin-
`<https://ekgf.org/cars/for-sale#gremlin2>`
Pro(s): Easy to implement

Con(s): Chance of duplicates and collisions high

Scenario / Use Case- Any and all attributes can be used to make the URI somewhat unique for identification and can be used for the URI

URL Encoded name concatenated to the URI .

<https://ekgf.org/cars/for-sale#gremlinUsed%20Vintage%20Gremlin%20%237>
<http://purl.org/goodrelations/v1#name> "Used Vintage Gremlin #7"

Pro(s): Easy to implement
Con(s): Chance of duplicates and collisions moderate, uses more details from the attributes. Discouraged, considered an anti-pattern for designing ontologies.

Scenario / Use Case- Use a unique identifier from an external service in the URI .

<https://ekgf.org/cars/for-sale#gremlin> Could have a Vehicle Identification Number added to the URI, the organization responsible for VINs guarantees uniqueness <https://ekgf.org/cars/for-sale#gremlinA3A465H399999>

Pro(s): Easy to implement
Con(s): Chance of duplicates and collisions moderate, uses more details from the attributes

Scenario / Use Case- We can promote blank nodes to more explicit local resources in the form urn: and/or _id:

Perhaps we adopt urn:org:method:id, method could be base64, sha256, or others which would indicate the type. could also look like:

create a blank node for VIN of the vehicle _:A3A465H399999

<urn:ekgf:vin:_:A3A465H399999>

Pro(s): Gives a more normalized approach for blank nodes to be used with other systems. Or not-well defined canonical scheme.
Con(s): Debatable utility if you have a namespace already, if possible use a prefix instead.

Scenario / Use Case- We can use an iso8601 date timestamp for a unique identifier for the URI .

iso8601 generated date/time
2021-11-22T03:28:40+0000
URL Encoded
2021-11-22T03%3A28%3A40%2B0000
<https://ekgf.org/cars/for-sale#gremlin2021-11-22T03%3A28%3A40%2B0000>

Pro(s): Simple to implement with a timestamp, quite reasonable
Con(s): There is still a very small chance of duplicates.

Scenario / Use Case- We can use a UUID for a unique identifier for the URI .

Generated the following UUID : e37b55e6-7447-4427-bcd8-9dae64750a1d , concat to the URI .

<https://ekgf.org/cars/for-sale#gremline37b55e6-7447-4427-bcd8-9dae64750a1d>

Pro(s): Simple to implement with a uuid, quite reasonable
Con(s): Not repeatable, will need to store as each generated identifier is unique

Scenario / Use Case- We can compute SHA 256 for which we can use as a basis of identity from attributes of an object. SHA hash values have the advantage of being reproducible vs UUIDs/iso8601 dates.

In this example we compute SHA 256 from the name "Used Vintage Gremlin #7"

<https://ekgf.org/cars/for-sale#gremlin>
<http://purl.org/goodrelations/v1#name> "Used Vintage Gremlin #7"

URIs with a SHA 256 will need an ALPHA character leading it as an URI cannot lead with a NUMERIC

Attributes of an object can be used to make the object URI unique by using a SHA 256.
<https://ekgf.org/cars/for-sale#gremline1752b9574895a962d4bfa554b98296d3acc06fdd5871f75f1c902ed3d5c13c3>
<http://purl.org/goodrelations/v1#name> "Used Vintage Gremlin #7"

SHA256 can be regenerated in a one way manner vs UUID which cannot. UUIDs may need to be managed once generated.

Pro(s): Simple to implement like a uuid as SPARQL and various languages have support, quite reasonable, also repeatable one way
Con(s): Some may find the length of the identifier an eyesore, that said the length is completely reasonable and safe for URIs if a bit challenging to type

Scenario / Use Case- Use an AES (or other methods such as ChaCha) encrypted querystring for additional security.

It may be desirable to encrypt a query string / part of a URI to prevent tampering and obfuscating parameters. Note that this should not be relied on for access control, so username and password are perhaps not good candidates for encryption- instead http headers should be used per OWASP .

If we take the VIN A3A465H399999 from the gremlin and create a SHA256 from the VIN .

SHA 256 of A3A465H399999

27d68368497b314c6bae0fa8ae61fd2c78406c01f3202c0bd73a63f1065e668b

We then could encrypt the SHA 256 using a key !!!!!!!ekgfknowledgegraph!!!!!! mode ECB . Base64 encoded results are agMNw71rsJcJOrMdTIVKEQqfwqmU56HtzZozR2DYjrvsyr9+bJXMfVZk2x0hqwz55njtWLMmzcwPqYDz6ovJW1tODpnWD60srd9046E/FWE= Which could then be incorporated as part of the URI

<https://ekgf.org/cars/for-sale#gremlinagMNw71rsJcJOrMdTIVKEQqfwqmU56HtzZozR2DYjrvsyr9+bJXMfVZk2x0hqwz55njtWLMmzcwPqYDz6ovJW1tODpnWD60srd9046E/FWE=>

Decrypted Base64 would yield

MjdkNjgzNjg0OTdiMzE0YzZiYWUwZmE4YWU2MWZkMmM3ODQwNmMwMWYzMjAyYzBiZDczYTYzZjEwNjVlNjY4Yg==

Converted to plain text from Base 64 would yield the SHA 256 of the VIN
27d68368497b314c6bae0fa8ae61fd2c78406c01f3202c0bd73a63f1065e668b
Querystring parameters besides an identifier, can also be handled by AES encryption

Pro(s): Provides extra encryption for IRI/URI
Con(s): Some may find the length of the identifier an eyesore, and the AES encryption excessive and perhaps if such encryption is required the information shouldn't be in an URI

Scenario / Use Case- Use a Decentralized Identifier / DID

This could also include distributed identifier services from the w3c, the below is non registered but taken from a sha256 of a VIN,then encrypted using AES256 with key !!!!!!!ekgfknowledgegraph!!!!!!! using ECB mode from the previous encryption example. The format of a DID is :

did:method:key
did:ekgf:agMNw71rsJcJOrMdTIVKEQqfwqmU56HtzZozR2DYjrvsyr9+bJXMfVZk2x0hqwz55njtWLMmzcwPqYDz6ovJW1tODpnWD60srd9046E/FWE=

Pro(s): Leverages an external organization to create the identifier, emerging as a standard.
Con(s): Leverages an external organization to create the identifier, specs and software may need to be written for internal usage.

D.2.11 Knowledge Graph Storage and Retrieval

Description

Semantic data storage managed by the EKG/Platform

D.2.12 Knowledge Graph Federation

Description

Model-driven federated interoperable and distributed data platforms

An Enterprise Knowledge Graph (EKG) can be seen by end-users as "one thing", a "holistic" collection of all connected data, similar to the web.

However, as also specified in principle 'Distributed' on page 78, the EKG is distributed by nature, assuming that it is not realistic in very large organizations or even eco-systems to have only one physical implementation of a fully centralized EKG. That means that different parts of the EKG are served by different installations or deployments owned and controlled by different parts of the organization or even other organizations.

Each deployment can be configured to connect to any number of ”backend” data sources (or destinations/sinks), some of which can be real triple stores (aka RDF Databases, quad stores or semantic graph databases) and some of which can be relational databases, key/value stores or any other database technologies. Or even just services with APIs that are used to get or store data.

At higher levels of EKG platform maturity, all access to the EKG is provided via this service layer — generally called the EKG/Platform — that takes care of federation of any request to any backend data source using any technology.

All that technology is however hidden for the user. In that sense an EKG/Platform is just a SOA layer[1]. However, it is a fully model-driven SOA layer and one that works with all other known deployments of the EKG/Platform.

The federation facilities provided by the EKG/Platform are leveraging the principles of the linked data standard$_{\text{linked-data}}$ as originally defined by Sir Tim Berners-Lee in 2005. However, the original linked data standard does not provide many of the facilities that are required for mission-critical enterprise use cases such as model-driven entitlement enforcement, automatic selection of the right version of the truth for the given context and so forth.

Various query protocols have built-in federation facilties, for instance the SPARQL query language has a facility — via the `SERVICE` keyword — to federate a query across multiple remote endpoints and the GraphQL query language can combine APIs of multiple remote systems. Regardless of the query protocol used, federation provides a link that abstracts the underlying system in a way that seamlessly ties sources together.

An EKG/Platform provides capabilities to federate queries on the backend.

The EKG/Platform is discoverable by other services, EKGs or browsers a client can also federate remote knowledge graphs not only in server backends, but even in browser based faceted implementations.

An extension of federation is virtualization, which will be discussed in further section. Virtualization provides linked data transformation from a source not designed for knowledge graph in a materialized or ad-hoc manner using a mapping facility.

D.2.13 Knowledge Graph Virtualization

Description

Model-driven views on data of any source leaving it in-place

(Work in progress)

Virtualization provides knowledge graphs the capability to source data not designed for linked data concepts like relational data sources.

It is one of the many techniques that are available in the arsenal of EKG engineers to get data from a given source to be ”connected” to all other data in the EKG.

Generally these are the options at a high level:

- The EKG itself is the authoritative source of a given dataset.
  - The dataset is complex
    - Use a triple store (no virtualization needed)
  - The dataset has a very large volume (> hundreds of billions of facts)
    - Use a specialized database type and apply virtualization.
- The EKG is not the authoritative source of a given dataset.
  - The data is relatively clean and there is either a real-time requirement or a massive volume of data.
    - Virtualization — as provided by various vendors like Stardogand Ontotext GraphDB — may be a good option because it gives an easy path from SPARQL to whichever backend database, translating SPARQL ”on-the-fly” to SQL or other query languages.
  - None of the above:

[1] See `https://en.wikipedia.org/wiki/Service-oriented_architecture`

* Process all data from the given source in one batch Extract Transform Load (ETL) pipeline, store it in a triple store controlled by the EKG/Platform and serve it from there. No virtualization needed, therefore more opportunities to enhance quality in the pipeline (since virtualization often comes at the price of being less flexible in terms of available options to enhance data quality).

By the use of RDF mapping languages, relational data is mapped to semantic knowledge graphs leveraging tools that implement RML, R2RML, OBDA, YARRRML, D2RQ and many vendor specifications.

Lossless conversion to an RDF data model can be achieved by defining terms to relational schema definitions of columns, type, and tables in mapping definitions. Depending on the tool used to perform the virtualization, data can be materialized into RDF files and/or ad-hoc on demand. A knowledge graph virtualization tool that operates on relational data will execute a SQL query and convert the results to RDF. If the results of the virtualization is materialized, the query executed and the results stored as RDF files. If the virtualization is performed ad-hoc and not stored, the query will be excused to the source system each time the query is performed. If there is a large resultset expected from virtualization, it usually more performant to materialize the data and load the RDF into a knowledge graph. If a small detail based result set that is susceptible to change is expected from the virtualization, then ad-hoc is desirable. Virtualization also can be done on heterogeneous data outside of the knowledge graph, before an RDF mapping language is applied. Data virtualization can be used to allow multiple data sources to be combined and queried as a single data source. External data virtualization can be used to get past limitations for example with R2RML which defines a single data source. Many relational data and knowledge graph virtualization tools also support Spark. Mapping can be very prescriptive to an ontology from a source system, or less. The less prescriptive approach is to allow data simply to be converted to an RDF model, later to be modified by SPARQL queries. The more prescriptive approach requires more diligence in model mapping.

Chapter D.3

User Interface

The User Interface component has the following capabilities:

D.3.1 Search

Description

Standards and tools for finding relevant data across the enterprise.

D.3.2 Collaboration

Description

...

D.3.3 Multilingual Support

Description

...

D.3.4 Content Editing

Description

...

We welcome your input here.

D.3.5 Reporting

Description

...

D.3.6 Visualization

Description

...

Appendices

Appendix A

Manifesto

As a reference, we include the manifesto chapter of the EKG/Manifesto document here as an appendix.

1 Ideals

Should we specify higher level ”ideals” to better explain what drives us? This is a section of ”reasons why” in non-data / non-tech terms. This section is very early days and needs much more work.

We welcome your input here.

1. Much higher levels of transparency, sustainability, fairness and accountability are going to be required at all levels in any organization
2. Human Capital needs to be known, valued, leveraged and optimized
3. Data Capital needs to be known, valued, leveraged and optimized
4. Increasing competitiveness depends more and more on having the highest quality and depth of data, information and knowledge
5. One ”censored”, biased, Single Version of (the) Truth (SVOT) is no longer good enough for many — if not most — use cases in most domains
6. The world becomes more and more polarised due to ”information bubbles” that many people are not escaping from, more depth and connectedness is needed
7. The world becomes more and more complex and harder to understand, a holistic view around every given topic, showing all viewpoints, would help people to understand and make better decisions

2 All data connected

1. All data will be connected.
    - Information, Knowledge, Meaning: it’s all data.
    - Knowledge & Meaning will be captured as machine-readable executable models.
2. All data will be made available anywhere — within entitlement limits — at any time to any device, node or edge.
3. We embrace the *Open World*[1] and deal with the realities of “multiple versions of the truth (MVOT)”.
4. We combine the digital footprint of activities along with a digital representation of information and knowledge, from which an EKG emerges.
5. All your connected data is an EKG.
6. An EKG is connections of Knowledge Graphs across an Enterprise and beyond.
7. We encourage the use and widespread proliferation of EKG identifiers

[1] Does not mean all data is open to everyone

3 Standards

1. EKGs are based on standards and therefore interoperable across boundaries.
2. There are many different types of standards that an EKG needs to be able to deal with.
3. Standards are described as machine-readable models — i.e. ontologies — that EKG/Platforms can execute, interpret or enforce.

4 Data Markets

1. Any data source will be turned into a data publisher — or supplier — of one or more SDDs.
    - Any data sink will be turned into a data consumer — of one or more datasets.
2. Data suppliers and consumers will find each other via a data market using a standard ”lingua franca” for the data itself, its meaning, all its associated policies and metadata and especially also its use cases as executable models.
3. The data market manages the information supply chains between all the various suppliers and consumers.
4. The global data market will consist of many other more specific data markets e.g. per industry or per enterprise.
5. An EKG is the combination of one or more data markets and the deployment of its use cases.

5 Open World

1. For any given ”Thing” — i.e. an object — there may be many representations in many datasets.
2. An object’s representations may be different in shape, meaning, timeliness, relevance and quality — i.e. any given representation of information about a given object may represent a different version of the truth.
3. All representations of any given object shall be linked via shared identifiers.
    - Identifiers shall be meaningless, opaque, web-resolvable and universally unique. See principle 1.
    - An object can have multiple identifiers.
4. Any given object consists of 1 or more Data Points.
5. A Data Point represents a logical property of a given object.
    - The identifier for a Data Point is the identifier of the object it belongs to plus at least one identifier of the axiom that describes its meaning (which is also an object).
6. Data Points for the same object can exist in many different datasets.
    - with potentially multiple versions of the truth in terms of meaning, timeliness, relevance and quality.
7. Any representation for any given Data Point of any data source shall be made available to any device, node or edge in the network within legal, policy and entitlement limits in real-time.
8. Every ”object” that is represented as data in whichever dataset anywhere, shall have an identifier that is universally unique, permanent, meaningless or opaque and therefore shareable, resolvable through the HTTP protocol. See 1 Identity.

Work in progress notes: that add more explanation to the above:

- Data will be considered explained when its usage has no misconceptions nor ambiguities.
- The word “data” has a lot of different notions associated with it. We have these statements above, like “all data is connected”, “…data will be made available anywhere…”, which could reinforce a particular notion that data is something that “exists” all around us, like a digital footprint of activities. On top of these statements on “data”, the EKG is introduced as a combination of one or more data markets, which could further reinforce that ”EKG is connected data”. If that’s the notion that the manifesto wants to declare, it may appeal strongly to some sections more than others, like those dealing with consolidation, analysis, reporting, verification etc. If the manifesto’s intent is broader (here, I am very conscious of Carl’s note that it should stay away from any hubris!) it would help if there is a way to declare that information and knowledge is also data. And it is when we combine the digital footprint of activities along with a digital representation of information and knowledge, that a Knowledge Graph emerges. And on top of that notion, an Enterprise Knowledge Graph is connections of Knowledge Graphs across an Enterprise. The

intent of the above is to appeal broadly to different sections of an Enterprise, accommodating different norms and notions (again referring to Carl's insightful comment on Norms and Notions)

Appendix B

Principles

As a reference, we include the principles chapter of the EKG/Manifesto document here as an appendix.

As a reference, we include the principles chapter of the EKG/Manifesto document here as an appendix.

These principles are intended to provide guidelines for the development and deployment of an Enterprise Knowledge Graph (EKG). The principles emphasize shared meaning and content reuse that are the cornerstone of operating in complex and interconnected environments.

1 Identity

Any object in the EKG is identified with at least one universally unique, opaque, permanent, non-reassignable and web-resolvable identifier in the form of an IRI (Internationalised Resource Identifier) for the EKG — i.e. an "Enterprise Knowledge Graph IRI (EKG/IRI)".

The EKG/IRI identifier is permanent, will be proliferated across the company's universe (including ecosystem), and will be used for the expression of facts about the object including relationships between objects.

Additional non-EKG identifiers may also be assigned, and they may be human-readable, "external" to the organization's EKG and be transient and reassignable.

Resolving an identifier can be done in three ways:

1. using it in a transaction — i.e. a query or update statement — submitted or routed via an internet protocol to a "lookup service" that translates one or more given "features" of an object to an EKG/IRI.
2. constructing it via a standardized policy from key components and applying a hash and optionally signing it — where the object represented by the EKG/IRI may or may not already exist.
3. constructing it by giving the object an EKG/IRI based on a random number in case the EKG is the authoritative source for the given object.

Rationale While the semantic web technologies — like Resource Description Framework (RDF) — generally allow for many and varied "Internationalised Resource Identifiers (IRIs)", and this is still encouraged when integrating systems, there is benefit in being able to rely on one canonical and unchanging one, which can for example make the mapping of identities a many-to-one rather than a many-to-many task.

In addition to that, to enhance the likelihood that various EKGs — across departments, organizations or ecosystems — can interoperate easily with each other, the use of standardized EKG/IRIs needs to be encouraged since various EKGs can come to the same identifiers independently, drastically increasing the number of links across EKGs.

Implications

- There should be a mapping or service to resolve other names, keys or IRIs to the EKG/IRI.
- Since it is immutable, the EKG/IRI will have to be *opaque* i.e. not be a human-readable since even human names, company names, customer numbers, Social Security Numbers can change over time.

- Objects that already have a well-established RDF-compliant and Linked Data compliant identifier may not necessarily need an additional EKG/IRI. In fact, they may already have one that is external to the company's EKG. It is in many cases recommended to even give those well-established objects from well-established external datasets a company EKG/IRI. Examples of such objects are Legal Entity Identifiers (LEIs) and Financial Instrument Global Identifiers (FIGIs). The EKGF will maintain a list of these for convenience.
- The use of multiple identities generally means that an EKG should *not* use the Unique Name Assumption (UNA) (where use of a different identifier would imply a different object). However the UNA *would* apply specifically to the EKG/IRI, and this should be ensured by any service.

2 Meaning

The meaning of every Data Point must be directly resolvable to a machine-readable definition in verifiable formal logic.

A Data Point combines an object — using its EKG/IRI as its identifier — with the value of a *property* in some context. Hence data is expressed at its most granular level for both data at rest and data in motion.

The property itself is always an IRI, often called "predicate-IRI", that refers to an object[1] that represents "the meaning" of the given Data Point. This object has its own identity and is defined through further properties based on logic that allow information to be rigorously combined, queried and inferred. These properties that define properties — also called "axioms" — are standardized by means of the Web Ontology Language (OWL2) by the World Wide Web Consortium (W3C) and are grouped into *"OWL ontologies"* for management purposes.

Rationale Expressing data at a granular level allows ultimate flexibility for it to be sliced, diced, combined and aggregated. This capability to combine and infer information is further enhanced by the use of property definitions built on logic. Having the properties themselves be objects that can be looked up means that all data is self-defining and carries its meaning with it. Since the information is self-defining there is no fixed schema for the EKG as a whole and it can non-disruptively incorporate additional knowledge.

Implications Some further discovery, with subject matter experts and creators of source systems, is often needed to truly understand what a given set of data really means and what can be inferred from it. In other words you cannot rely on the name of a column in a spreadsheet. A deceptively simple column name such as "number of European customers" leaves open the meaning of "European" and "customer" and timing (when does one start and stop being a customer?). And different sources could have different interpretations of that same name. The benefit is consistency, accuracy and the ability to make sound business decisions.

Advanced at higher levels of EKG/Platform maturity the term Data Point may in fact become a more complex data structure that is used "on the wire" that represents the Data Point at a more "holistic" level, supporting "multiple versions of the truth (MVOT)". Since an EKG supports many datasets that have overlapping information coming from multiple sources, there could be:

1. multiple EKG identifiers (EKG/IRIs) for the same object
    - One object can have multiple identifiers that can be linked together [2] and therefore be rightfully addressable with any of these identifiers.
2. multiple definitions of meaning
    - One property of an object can have multiple definitions of meaning, for instance "legal name" can be defined in multiple ontologies and be semantically equivalent [3] or one property can be defined as a subproperty of another but broader semantic definition [4].
3. multiple equal or different values coming from multiple sources
4. multiple versions over time of those values (temporality)

For each of these four "axes" — identity, meaning, source, temporality — you could have multiple options to choose from even while logically, from a user perspective, it's the same data point. Advanced client applications, services or AIs can use

[1] the official term in the RDF standard for this object is "resource"
[2] for instance via owl:sameAs i.e. "Individual Equality"
[3] See Equivalent Object Properties or Equivalent Data Properties
[4] See Data Subproperties

these Data Points to perform last-minute "at the edge" computations around finding the right value from the right timeline and source with the right quality for the given context.

3 Distributed

Data Points for any object may be stored in many physical stores. Any access point provides connectivity to all EKG content regardless of where it resides.

The physical stores could include traditional databases as well as varied Platforms specifically designed for hosting EKGs. And some which may be hosted outside the enterprise including by information services or governments. All seamlessly accessed using W3C (internet-based) protocols which also allow for browser or API *linked data* traversal and query.

Rationale Federating different systems allows the knowledge available to an enterprise to be accessed as a whole. It allows best of breed systems to be used for specific purposes, and evolved over time without disrupting the EKG as a whole. It allows scalability through adding new hardware, swapping out systems, and optimizing access through moving data closer to where it needs to be processed.

Implications The potential for a loose and evolving nature does mean some degree of monitoring of access patterns and performance; and service level agreements for vital information access.

4 Open World

Information can vary over time, come from many internal and external sources, and be based on different identifiers and models. These *multiple versions of the truth* need to be reconciled on access by context.

Different sources might not always be consistent about the same Data Point. And that may be legitimate for various business, geographical, privacy, legal or timing reasons. Rather than try to force everything into a single overruling set of facts, an EKG allows them to coexist, with the access context used to make coherent selections which are consistent within themselves. In order to make those selections multiple people and systems must have transparent access to all facts (including source, identity, meaning and value-at-time information) about all objects. Machine-executable business rules may be used at query time to join instances of data and e stablish quality rankings.

Rationale This approach allows the EKG to represent the sometimes messy reality, which encompasses not only a variety of different orgarnizations and systems within the enterprise, but also external sources.

Implications Attention needs to be paid to maintaining sufficient context with each data set, and considering what is needed for each data usage use case or context, including quality.

5 Self-describing

An EKG is composed of a set of *self-describing datasets* that provide information about lineage, provenance, pedigree, maturity, quality, licensing and governance.

The properties in each data point are linked to their definition so the meaning is not in doubt. A Dataset definition supplements this with management information such as its pedigree (how/when was it derived/sourced?) and its provenance (where/who did it come from?). This applies whether the information is maintained in the EKG itself or accessed/loaded from existing enterprise systems (data at rest); or received as data streams/messages (data in motion).

Rationale This information is essential for data selection, enforcing policy and management of the ecosystem as a whole. As well as being essential for management, the definitions taken together comprise a knowledge *catalog* for the content of the EKG.

Implications The information needs to be maintained and made available on an ongoing basis. It also needs to be sought out for external sourced data, whether accessed in place or loaded into an EKG platform.

6 Measurement

The quality and characteristics of the managed knowledge must be measurable and measured. Measurement criteria are used to designate fitness-for-defined-purpose and must be actionable.

Quality comprises multiple dimensions including accuracy, timeliness, completeness, relevance, conformity, integrity. EKG technology allows greater power and flexibility for expressing quality rules and metrics, but existing data quality tooling could also be used.

Rationale This information is needed to allow the right information to be used for each use case.

Implication The information needs to be maintained and made available on an ongoing basis.

7 Business Orientation

All information in the EKG, and associated artifacts, are linked to defined and prioritized *use cases*. Nothing in the EKG exists without a known business justification and purpose.

An EKG use case encompasses a business narrative and outcome expressed in business terms and links to relevant ontologies and Datasets.

Importantly, the use cases for an EKG can make use of lower level use cases, thus forming a *use case tree*, though it is not a strict hierarchy since common use cases can be reusable. At an implementation level a use case may be associated with user stories that form the points of interctaion with end users or client systems. The vision is that a fully fleshed out and implemented use case can be deployed as a reusable component.

Rationale Use cases anchor the EKG to real business needs, and allow it to evolve incrementally while delivering business value. Without this, the tendency is often to focus on information modeling for its own sake, without a focus or rationale for what to include or exclude. The use cases provide the basis for associating users with EKG functionality and provide the context for information selection.

Implications The use cases themselves need to be developed and managed as part of the EKG development method, and ideally as part of the EKG itself.

8 Control

Entitlement, privacy and business policies will be modeled in the EKG and automatically executed, enforced and audited at the Data Point level.

The EKG can use enterprise and organization knowledge to express access not only in terms of access control lists, but in terms of business rules, policies, logic and information content.

Rationale Use of the EKG itself to control and enforce access allows more power and conciseness of policy expression and execution while linking to existing enterprise directories.

Implications Appropriate enterprise directories should be integarted in the EKG. It can take some thought to design what the policies should be at the business level.

9 Ecosystem

An enterprise will use a heterogenous set of technologies and data sources which will be incorporated into the EKG over time. All data entering the ecosystem are subject to service level agreements.

A true EKG is a federation of multiple datasources and systems both within and external to the enterprise; seamlessly knitted together with standard protocols and APIs. An EKG needs to be managed and evolved as a fairly complex system of systems. To provide the flexibility and agility needed, its management needs to be automated, and linked to development processes, through the discipline of *data operations*.

Rationale As an ecosystem, an EKG can non-disruptively evolve over time in technology, scalability and information and business needs addressed.

Implications The management of the EKG needs to be planned for and resourced. It's important to coordinate with the owners of existing systems being used to provide capability under the EKG umbrella.

10 Standards

Data, ontologies, definitions and business rules should be compliant with open standards and transparent governance procedures as defined by recognized standards bodies.

Where necessary, EKGF will work with relevant standards bodies or projects to propose new standards or enhance existing ones.

Rationale It's important for enterprises to have trust in the quality, interoperability and stability of the structures and interfaces used in such a strategic investment as an EKG.

It provides freedom of choice in being able to mix and match different technologies and models, either in a federated EKG or over time. That includes standards being able to respond to new advances in both technology and techniques.

Implications It can sometimes take longer to work with industry partners to reach the consensus needed for standardization, and to use change control/governance procedures. Release cycles may be phased to avoid constantly changing interfaces. In order to achieve agility some work products may be deployed on standards that are have a *provisional* status — these need to be carefully identified.

Appendix C

EKG Foundation

This document — the Maturity Model for the Enterprise Knowledge Graph (EKG/MM) — is one of the key artifacts produced by the Enterprise Knowledge Graph Foundation (EKGF).

Vision and Expectation Management It captures our thinking around the vision of EKG, how it affects your business and what can be expected of it at certain levels of capability maturity.

Non-functional requirements The EKG/MM is therefore seen as the overall framework for all "non-functional requirements (NFRs)" — requirements — clustered per capability — that are not directly related to any given "use case".

Use Case Tree For instance, if it is your company's strategy[1] to be able to have real-time risk insights at all levels of your organization then we can separate the specific functional requirements of that very large use case from all the NFRs that would come into place around your data management and governance practices, the state of your technology landscape, the culture of your organization and so forth. In this example, the use case *Real-time risk management* — as the top-level strategic use case — would be broken down into various *sub-use cases* resulting in a so-called *"Use Case Tree (UCT)"*. Each node in that tree represents some functionality — a capability so you will. How does this relate to the EKG/MM? Well, per use case in that tree structure you can determine how realistic it is to actually be able to implement that use case in relation to the level of maturity of the organization as a whole and more specifically in relation to any of the capabilities that are being covered in this document.

Use Cases Replace Silos Another major point to make about the Use Case Tree (UCT) is that it is the backbone component-structure of your EKG. In a way you could see them as the EKG's equivalent of your current siloed application- and data-landscape. Rather than building a new silo everytime for every new use case, you would "build" new use cases in the EKG. Use cases are *logical silos* in that sense. In a data-centric EKG world, there are no more independent applications that are run as separate systems. Just this particular change alone has major implications on how things are organized. The EKG/MM covers these — and many other — implications.

Capabilities as Use Cases The EKG/MM is about the generic capabilities and how they may evolve over time and how some of them are required to be available at a certain level of maturity. That being said, almost every capability that we describe in this document is in itself also a use case. Many if not most of these capabilities will more and more depend on high quality detailed and "holistic" data.

Other Artifacts

Manifesto & Principles for the Enterprise Knowledge Graph (EKG/Manifesto) The EKG/Manifesto is a combination of long term vision statements — the "Why" — combined with the underlying principles for EKG — the "How".

Enterprise Knowledge Graph Method (EKG/Method) The assumption is that each organization in a data-intensive industry will end up owning their own EKG — possibly shared across organizations — and therefore requires oversight by a group of people — however organized — that we call the CoE for the EKG.

EKG/Method describes all the practices of the CoE for the EKG, from inception to production.

[1] See 'What is Strategy?' on page 84

EKG/Catalog The EKG/Catalog describes all the common use cases (and some of their "sub-use cases") for EKG. The goal is to also publish these use cases on the ekgf.org website under the so-called "use case portal" and make them available for download as modules, basically the same kind of ecosystem of reusable components as any other successful technology-stack has. This is comparable to what "Maven Central" is for the Java world or what "npmjs.com" is for the JavaScript world.

Use Case Central One of the primary goals that the Foundation is working towards iś–ás briefly mentioned abové–t́o publish use cases on the ekgf.org website, initially just as a searchable catalog of use cases, just text basically, but soon after as a graph, as part of an EKG. Some members of the Foundation — in particular agnos.ai and eccenca GmbH — already have EKG infrastructure running in production that allows them to run "no-code" or "Low-Code" use cases as full transactional applications. The idea is to standardize this, to publish use cases as "reusable components", with all their dependencies, ready for immediate deployment, just like one would do in the Java/JVM world or in the JavaScript world or with any other successful technology stack.

Portals The Foundation is still in its start-up phase. Plans are to invest in the development of a more advanced website that consists of various "portals" focusing on serving a specific need for a specific audience:

Use Cases
- Short term:
  - Audience: Business Executive, Vendor
    - Understand the business problems that can, and have, been addressed
    - Use a framework for structuring projects
    - Get directed to relevant reuse points in other portals (e.g. for ontologies)
- Long term:
  - Audience: Executive, Vendor, Knowledge Graph Engineer
    - Identify, select, download & provision production-ready use cases

Best Practices
- Audience: Project Lead, Consultant, any role in the CoE for the EKG
  - Reduce the risk and cost associated with new EKG projects
- Audience: Vendor
  - understand how to develop, sell and deploy their products (and services) to maximize applicability and success

Software
- Audience: Architect, Vendor, Academic
  - Access reusable software curated for EKG purposes (may be hosted externally)
- Audience: Vendor, Academic
  - Position existing products or components to increase uptake

Ontologies
- Audience: Modeler
  - Assess ontologies curated as reusable for EKG use cases
  - Apply tooling to automate ontology development and measurement
- Audience: Academic
  - Access a set of ontologies for research, analysis, or extension

Datasets
- Audience: Analyst
  - Access reusable RDF resources curated for EKG purposes (may be hosted externally)
- Audience: Analyst, Vendor
  - Make use of data resources for demonstration and experimentation

Member Directory & Services
- Connect with people and parties with the knowledge to help (see also appendix Corporate Members)
- Promote your knowledge and skills

Persona Stories The Foundation recognizes different types of "personas" that it needs to serve:

<as a>	<I want to>
Business Executive	Realize the benefits and minimize the risks of EKG by having access to proven methods, best practices and a community of experts.
Vendor	Make it easy for organizations to understand and successfully adopt my product across multiple industry sectors.
Modeler	Have access to a proven and consistent set of deployable use cases, models & ontologies that can be used with my organization's EKG.

Technical Architect	Have access to components and interfaces with supporting technology architectures that I can assemble and deploy within my environment.
Consultant	Have access to — and contribute to — EKG-related best practices (e.g. EKG/Method & EKG/MM) as well as a community of potential customers and skilled associates.
Academic	The chance to make a meaningful and recognized contribution that builds upon frameworks to address pressing business needs.
EKG Engineer	be informed about best practices, role descriptions, education and certifications.

Appendix D

What is Strategy?

Competitive strategy is about standing out from competitors by offering a unique mix of value. Michael E. Porter defines strategy as; *"Strategy is the creation of a unique valuable position; involving a different set of activities"* [Porter].

Valuable positions emerge from three different sources. First, positioning can be based on producing a particular set of products or services. Second, positioning can also be achieved by targeting the needs of a particular customer segment. Third, the last basis for positioning is based on the way customers can be reached as the way to reach a particular segment might differ from other ones even though their needs are very similar. For example, customer could be geographically spread or could be served only through specific channels. Building a strategic position also involves trade-offs in terms of (1) image or reputation, (2) activities and resources, and (3) priorities.

Even though these trade-offs might seem undesirable at first sight, they are necessary for a sustainable strategic position according to Porter. These trade-offs reduce the risk of imitation and straddling. The former occurs when a competitor repositions itself to obtain the valuable strategic position of another company. The latter is far more common and occurs when a straddler seeks to obtain the advantages of a valuable strategic position while maintaining its existing position. However, an imitator or straddler may encounter many problems when trying to obtain a strategic position different from the one it already has. For example, a premium retailer may encounter many problems while trying to position itself as a low-price retailer as both positions require different priorities, resources and activities. This transition may even jeopardize the retailer's current strategic position and confuse its consumers by showing them two incongruent images. Therefore, Porter also defines strategy as; *"Strategy is making trade-offs in competing. The essence of strategy is choosing what not to do."*

Positioning determines not only what activities a company will carry out and how it will set up individual activities but also how activities relate to one another. The latter characteristic is called fit and the most valuable fit is strategy-specific because it enhances a position's uniqueness and amplifies trade-offs. Creating a fit between activities that is difficult to imitate also creates a sustainable strategic position. Finally, Michael E. Porter also defines strategy as; *"Strategy is creating fit among a company's activities. The success of a strategy depends on doing many things well — not just a few — and integrating among them."*

Appendix E

Business Operating Model

> *"The operating model is the necessary level of business process integration and standardization for delivering goods and services to customers. An operating model describes how a company wants to thrive and grow. By providing a more stable and actionable view of the company than strategy, the operating model drives the design of the foundation for execution."*
>
> Ross, Weill, and Robertson [RWR06]

Defining an operating model in an organization is very important because it enables the adoption of different strategic initiatives as a foundation for execution. Such an operating model represents and reinforces the strategy[1] to be supported. Therefore, the operating model — or the lack of it — has a profound impact on how a company implements business processes and IT.

An operating model has two dimensions which are (1) business process standardization and (2) integration:

1. According to Ross, Weil and Robertson (2006), standardization of business processes can be defined as; *"defining exactly how a process will be executed regardless of who is performing the process or where it is completed. Process standardization delivers efficiency and predictability across the company"*. Business process standardization can lead to large increases in throughput and efficiency. However, it comes at a price since the more business processes are standardized, the more difficult it is to tailor services and products for specific customer needs hampering innovation and business agility in the process.
2. Business process integration coordinates efforts and work across organizations through data. Sharing data can occur between processes to enable end-to-end transaction processing, or across processes to allow the company to present a single face to customers. Some of the benefits of integration are increased efficiency, coordination, transparency, and agility. However, integration entails a large amount of effort such as in end-to-end integration where companies need to develop standard data definitions and formats that will be shared and used across different business units and functions.

Companies adopt an operating model at the enterprise level but could also adopt different operating models at the division, business region, or other level. In order to define an operating model at any of these different levels, companies need to answer the following two question made by Ross, Weil and Robertson (2006):

1. *To what extent is the successful completion of one business unit's transactions dependent on the availability, accuracy, and timeliness of other business unit's data?*
2. *To what extent does the company benefit by having business units run their operations in the same way?*

The first question determines integration requirements while the second one covers the standardization requirements of an organizations. Depending on these requirements, there are four different possible operating models as depicted in table E.1.

Each one of these four operating model types is described below:

Diversification low standardization, low integration

[1] See 'What is Strategy?' on page 84

Business Process Integration	Business Process Standardization: Low	Business Process Standardization: High
High	**Coordination** • Shared customers, products or suppliers • Impact on other business unit transactions • Operationally unique business units or functions • Autonomous business management • Business unit control over process design • Shared customer/supplier data • Consensus process for design of IT infrastructure services • IT application decisions made in business units	**Unification** • Customers or suppliers may be local or global • Globally integrated business processes often with support of enterprise systems • Business units with similar or overlapping responsibilities • Centralized management often applying functional/process/business unit matrix • High level process owners design standardized processes • Centrally mandated database • IT decisions made centrally
Low	**Diversification** • Few if any shared customers or suppliers • Independent transactions • Operationally unique business units • Autonomous business management • Business unit control over process design • Few data standards across business units • Most IT decisions made within business units	**Replication** • Few if any shared customers or suppliers • Independent transactions are aggregated at high level • Operationally similar business units • Autonomous business unit leaders with limited direction over processes • Centralized control over business process design • Standardized data definitions, but data locally owned with some aggregation at corporate level • Centrally mandated IT services

(c)2005 MIT Sloan school of business

Table E.1: Operating Model Quadrant

Coordination low standardization, high integration

Replication high standardization, low integration

Unification high standardization, high integration

Diversification

Diversification applies to companies that have different units with few common products, services, customers, or ways of doing business.

Central management exerts relatively little control over business units that operate in a highly autonomous way offering their own products and services to their own customers.

The Diversification model may offer synergies from related, but not integrated, business units. In this context, business units may create demand for one another or increase the company's brand recognition which creates enterprise-wide value. Although there could be some synergies between business units, the success of companies with a Diversification model stem from the success of the individual business units and acquisitions of other related businesses.

Coordination

Business units in a Coordination company usually share one of the following: customers, products, suppliers, and partners. Some of the benefits of the Coordination model are integrated customer service, cross-selling and transparency across supply chain processes. Whereas key business processes are tightly integrated, business units have unique processes and capabilities.

In these companies, low cost is not the main driver as the main driver is delivering the best service and products to the customer while executing business processes in the most efficient way possible. Strong central management defines and prioritizes cooperation.

Through integration without a high degree of standardization across business units or functions, growth can be achieved by offering already existing products or services to customer segments in new markets. Additionally, growth can also be achieved by improving services to meet new, but related, customer demands.

Replication

The Replication model provides autonomy to business units but runs operations in a highly standardized way. The business units are not tightly integrated as they are not dependent on each other's transactions while they implement a set of highly standardized business processes that can be easily repeated in new business units. McDonald's and other franchise companies are examples of a Replication company. The advantage of this model is that it enables organizations to build new business units from scratch with relatively little effort.

Unification

Companies that operate as a highly optimized whole — around a highly standardized set of business processes — may benefit from the Unification model. Business units in these companies may have relatively very little autonomy and they best maximize efficiencies and customer services by using integrated data and driving variability out of business processes.

Unification companies typically have integrated supply chains with interdependencies between distributed business units. These business units share transaction data and standardized business processes. Therefore, these companies may benefit the most from enterprise-wide systems to support company standardization and integration requirements.

Management in these companies is highly centralized and plays an important role in driving out inefficiencies to foster growth through economies of scale by introducing new products. Since variability must be minimized in these companies, this model is best suited for companies that compete on price such as those that offer commodities where innovation or customization are not key.

Applying the operating model to attain growth

An operating model is the underlying logic of how an organization will enable and execute strategies. Each operating model entails its own opportunities for growth.

The need to tightly integrate business processes make acquisitions and mergers — for both the buy-side as well as the sell-side — more challenging as disparate data definitions need to be reconciled. However, the tight process integration of the Coordination and Unification models offers opportunities of organic growth through expansion into new markets or extensions of current product lines.

Process standardization, as in the Unification and Replication models, enables growth through a rip-an-replace approach to acquisitions. When an acquiring company wants to create a mirror image of itself out of an acquired company, it only has to replace the processes and systems of the acquired business with its own. However, both models do not offer much leverage when a company chooses to expand into operationally distinct lines of business as both models depend on leveraging already existing processes.

The Diversification model imposes fewer constraints on the organic growth of individual business units and fewer limits for business acquisitions. However, it does not offer the benefits of integration and standardization across business units. The opportunities for growth of each operating model are shown in table E.2.

Business Process Integration	Low Business Process Standardization	High Business Process Standardization
High	**Coordination** • **Organic:** stream of product innovations easily made available to existing customers using existing integrated channels. • **Acquisition:** can acquire new customers for existing products but must integrate data.	**Unification** • **Organic:** leverage economies of scale by introducing existing products/services in new markets; grow product line incrementally. • **Acquisition:** can acquire competitors to leverage existing foundation; must rip and replace infrastructure.
Low	**Diversification** • **Organic:** small business units may feed core business; company grows through business unit growth • **Acquisition:** unlimited opportunities; must ensure shareholder value	**Replication** • **Organic:** replicate best practices in markets; innovations extended globally • **Acquisition:** can acquire competitors to expand market reach; must rip and replace

Table E.2: Different operating models position companies for different types of growth

Business units of a company can also adopt different operating models to respond to conflicting demands. For instance, Diversification companies may benefit from allowing their own business units to adopt their own operating models as these business units are highly independent. An example of a Diversification company whose business units adopted different operating models is Johnson & Johnson. J&J's U.S. pharmaceutical group applies a Coordination operating model in which there is a single touchpoint with health-care practitioners while their subsidiary Janssen Pharmaceuticals implements a Replication operating model in Europe with highly standardized, low cost processes. This gives freedom to each business unit to implement a different operating model depending on its own objectives while maintaining a relatively simple operating model at a corporate level.

An operating model gives a company better guidance for developing IT and business process capabilities. It also serves as a stable foundation for strategic endeavours such as mergers and acquisitions. This foundation enables IT to be more proactive in identifying possible strategic opportunities. In order to define an operating model, management needs to define the role of business process standardization and integration. This also requires management to identify the company's key business processes that create a sustainable competitive advantage for the company. As a result, an operating model offers a company the possibility to create and possess reusable capabilities for long-term growth. In this context, an operating model could be seen as the main driver of strategy at a corporate or business level. In addition, an operating model plays a major role in defining the required architecture, practices, management thinking, policies, and processes as

they may be different for each operating model. In other words, an operating model could be a key driver in the design of separate organization units.

Appendix F

Enterprise Architecture

Once a company has defined one or several operating models depending on its business process integration and standardization requirements, it can then draw its enterprise architecture that reflects the company's key processes, systems, and data embedded at the company's core. More formally, an Enterprise Architecture is defined by Ross, Weil, and Robertson (p. 47, 2006) as "the organizing logic for business processes and IT infrastructure reflecting the integration and standardization requirements of the company's operating model."

It is essential to effective enterprise architecture to identify the processes, data, technologies, and customer interfaces that take the operating model from vision to reality as explained by Ross, Weil and Robertson (2006). These key elements are different for each of the four operating models.

Core Diagrams

A core diagram is a one-page picture of the high-level view of the processes, data, and technologies needed as a foundation for execution. The purpose of this core diagram is to facilitate discussions between IT and business managers to help clarify requirements for the company's foundation for execution and communicating the vision.

Although each core diagram can be unique to each organization, four common elements can be found almost in any core diagram, which are listed as follows:

- **Core business processes:** This small set of processes constitute the stable set of company-wide capabilities that the company needs to execute its operating model and respond to market opportunities.
- **Shared data driving core processes:** This is data shared across different end-to-end business processes by all the parties participating in these processes, such as business units or even suppliers.
- **Key linking and automation technologies:** There are different technologies that fall within this category. First, *middleware technologies* that enable integration of applications and access to shared data such as Enterprise Resource Planning (ERP) packages, data-warehouses, and knowledge graphs, are the first kind of these technologies. Moreover, portals providing standardized access to systems and data are the second type of these technologies. Finally, electronic interfaces facing stakeholders, such as suppliers, employees and partners could also appear on the company's core diagram.
- **Key customers:** These show the company's primary customer groups (e.g., channels or segments).

The four different types of core diagrams are discussed in the following subsections.

Enterprise architecture for a Unification Model

The core diagram for a Unification Model is shown in figure F.1. Integration and standardization of business processes is necessary to serve key customer types in a Unification model. The top of figure F.1 identifies the necessary and optional elements for business process designing, which are: 1. identifying key customers, 2. identifying and showing the key processes to be integrated and standardized, 3. identifying and showing the key data imperative to better integrate and standardize processes and serve customers, and 4. identifying and showing the key technologies that either automate or link processes (this optional as depicted by the dashed outline).

This core diagram provides a high-level overview of a company's enterprise architecture with highly standardized and integrated processes implementing a Unification operating model.

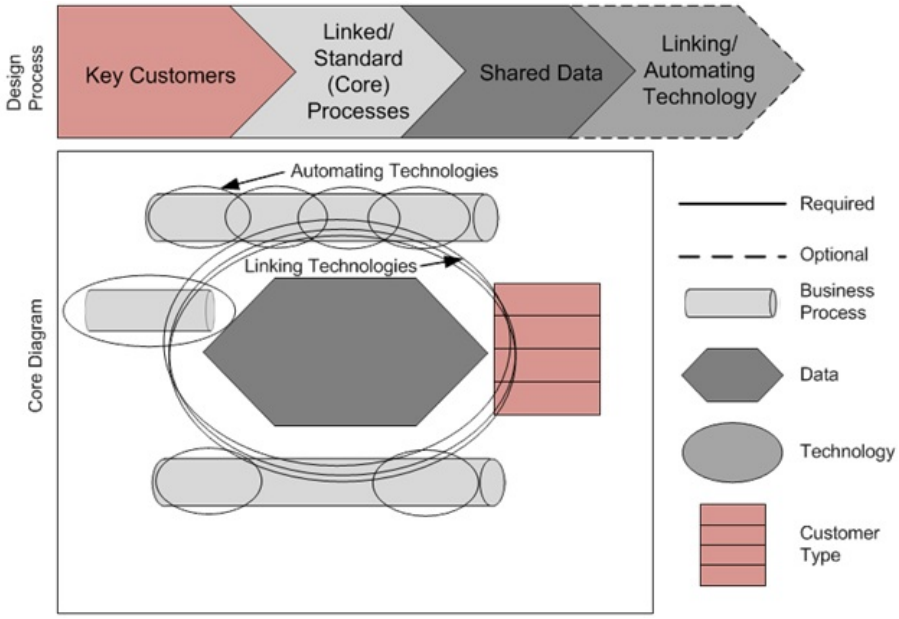


Figure F.1: Core Diagram Design Process4 (©2005 MIT Sloan Center for Information Systems Research and IMD)

Enterprise architecture for a Diversification Model

Since the diversification model entails the least amount of business process standardization and integration, it is the opposite of the unification model. In a unification model, each business unit functions more or less independently even though there can be opportunities for shared services across the company.

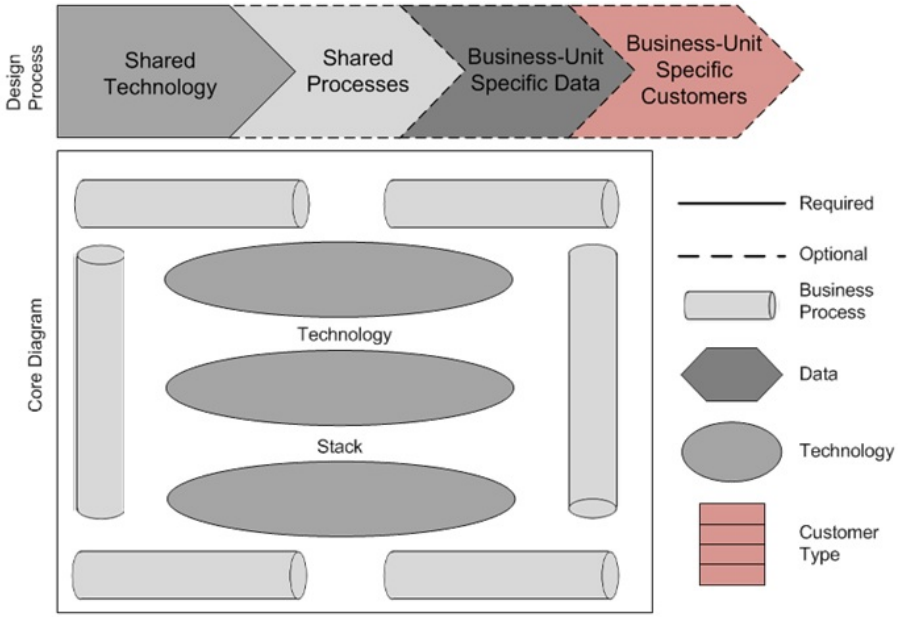


Figure F.2: Diversification Operating Model Core Diagram4 (©2005 MIT Sloan Center for Information Systems Research and IMD)

An extreme example of a Diversification model would be a total lack of an enterprise architecture in which there is no collaboration nor communication between business units. However, these business units could benefit from economies of scale by adopting a shared technology platform and are the key elements of the core diagram in figure F.2. Such shared technologies and services could take the form of data centers, telecommunications networks, help desks, offshore systems development and maintenance capability, centralized vendor negotiations, etc. If different business units may benefit from shared services and technologies, they could depict that in the core diagram of the company.

Additionally, if business units may also benefit from certain shared data, they could also add that to the company's core diagram.

Since shared platforms are the most important component of a core diagram of a company implementing a diversification model, start with the technologies that — when shared — could provide economies of scale, standardization, or other benefits.

The other remaining key elements — key customer types, shared processes, and data — should only be incorporated to the extent that they are common between business units and necessary for the operating model. For example, some data should be modelled in the core diagram as it may be shared at corporate level for financial reporting, risk management, and compliance across different business units.

Enterprise architecture for a Coordination Model

A company implementing a coordination model uses a high amount of business process integration to offer a service or product to each customer group. This high amount of integration results from sharing key data across business units to present a common face to the customer. Coordination allows companies to integrate many processes or products to the customer in a black-box fashion without forcing high levels of standardization.

The enterprise architecture core diagram for the Coordination operating model depicts the emphasis on integration and thus on shared data within a company. Commonly, the core diagram will also show important technologies used by stakeholders to access this shared data. Processes could also be shown in the core diagram if they need to be coordinated across business units. However, processes are often not depicted in the core diagram of a coordination model as they are mostly unique and can differ to a great extent across business units.

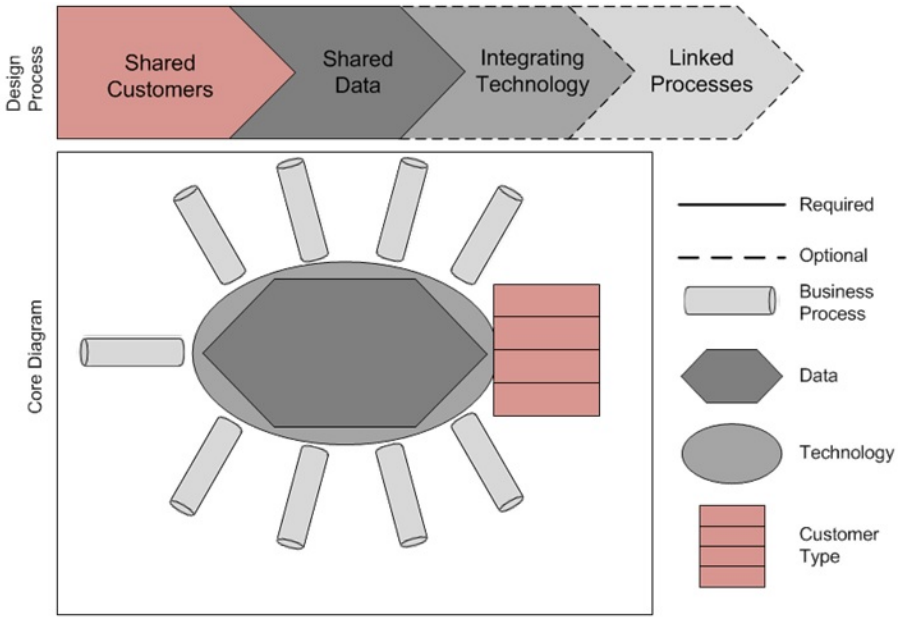


Figure F.3: Coordination Operating Model Core Diagram4 (©2005 MIT Sloan Center for Information Systems Research and IMD)

Enterprise architecture for a Replication Model

A Replication operating model is recommended when key processes need to be highly standardized but not integrated across the company by means of an automating technology.

Standardized processes is the key concept in Replication operating models. So, key standard processes and, often, the key technologies enabling those processes are the first things shown in the enterprise architecture of a company implementing a Replication operating model. In contrast to that, data rarely appears in the core diagram because business units in a Replication company barely share data. In order to drive up productivity, these companies could also automate key standardized business processes resulting in reusable business modules.

When drawing a Replication model core diagram (figure **??**), start with the key processes to be standardized and replicated across the business units. Then, identify the technologies automating those key processes. Next, look at the technologies

automating those key processes that could eventually be shared across the business units. However, since business units barely share data or have the same customer types as those decision are made locally, it is not necessary to add them to the core diagram.

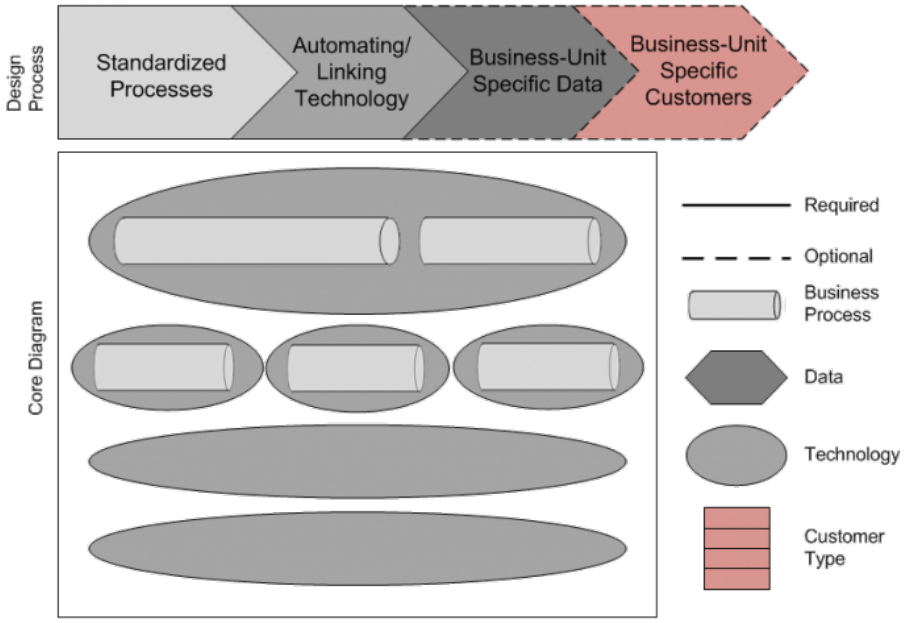


Figure F.4: Replication Operating Model Core Diagram4 (©2005 MIT Sloan Center for Information Systems Research and IMD)

The core diagram as a high-level overview of an Enterprise Architecture

The enterprise architecture core diagrams offered in the section are useful to have a high-level overview of the business process and IT requirements of a company's operating model. However, they do not offer the necessary detail to map out technical or process design requirements. A high-level enterprise architecture only offers a shared understanding of how a company will operate, but the synergies between people, process and technology necessary to implement that architecture demands shared understanding at a more detailed level.

Appendix G

EKGF Documents

Work in progress, create an overview of all major documents/deliverables that the EKGF is working on or going to work on.

EKG/Manifesto The EKG/Manifesto is a combination of long term vision statements — the ”Why” — combined with the underlying principles for EKG — the ”How”.

EKG/Method The assumption is that each organization in a data-intensive industry will end up owning their own EKG — possibly shared across organizations — and therefore requires oversight by a group of people — however organized — that we call the CoE for the EKG. EKG/Method describes all the practices of the CoE for the EKG, from inception to production.

EKG/Catalog The EKG/Catalog describes all the common use cases (and some of their ”sub-use cases”) for EKG. The goal is to also publish these use cases on the ekgf.org website under the so-called ”use case portal” and make them available for download as modules, basically the same kind of ecosystem of reusable components as any other successful technology-stack has. This is comparable to what ”Maven Central” is for the Java world or what ”npmjs.com” is for the JavaScript world.

Appendix H

Frequently Asked Questions

When does a KG get to be called an EKG?

Why are some but not all common horizontal activities selected?

TODO TODO TODO

Appendix I

How to contribute

1 Introduction

The EKGF is a non-profit member-driven organization that unites parties in the EKG space to define best practice and matures the marketplace for EKG adoption, including:

- semantic standards to solve the challenges of data management
- a method for accelerated EKG deployment

One of the activities of the EKGF is its work on the EKG/MM which establishes standard criteria for measuring and guiding progress on EKG adoption through asking practical questions of different involved stakeholders.

In its original form, it was a set of documents. In order to be able to scale this up to a larger community of collaborating teams, each contributing content to the EKG/MM, we decided to ”eat our own dogfood” and apply the same principles as we advocate for software development and data management such as ”work Agile”, ”use Git” and ”test everything”.

2 The Process

The collaboration process consists primarily of the following ”user stories”:

User Stories

- As a critic, I want to be able to **raise an issue**
- As a workgroup member, I want to be able to **plan issues**
- As a fixer, I want to be able to **fix issues**
- As a reviewer, I want to be able to **approve pull requests**

To summarize, we have 4 main things to do:

1. Raise (3)
2. Plan (4)
3. Fix (5)
4. Approve (6)

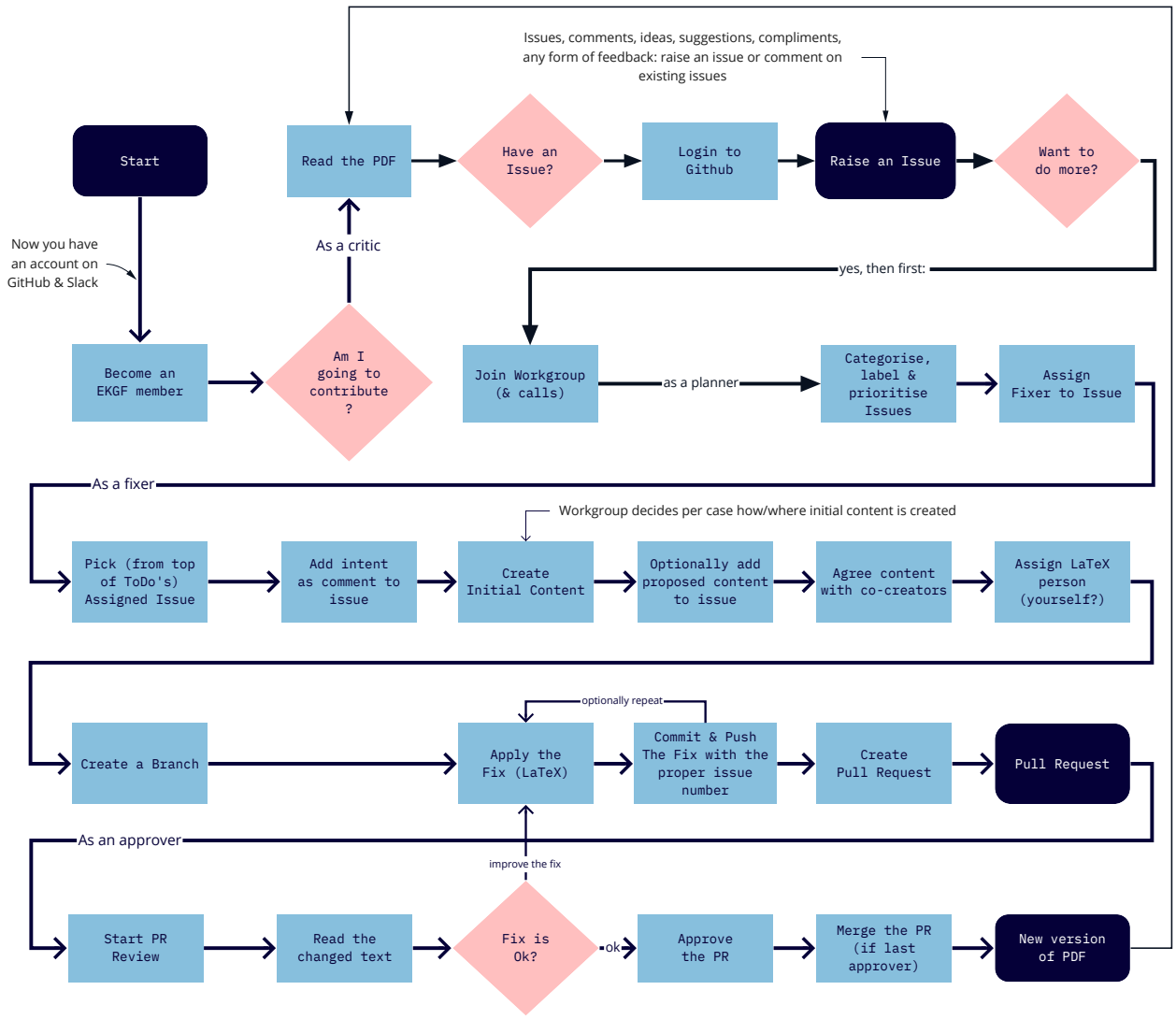
An issue can be anything, your feedback, ideas, suggestions, spelling mistakes, inconsistencies, anything that should be changed in the PDF document that you have just read.

3 Raise

User Stories

- As a critic, I want to be able to **raise an issue**

Figure I.1: Contribution Process Diagram



Anyone who reads the produced content and has an issue with that content is a ”critic”. Each critic should be able to raise that issue formally with minimal effort.

The audience of critics is hopefully very large, any reader of our content should be able to add their comments somewhere. We have to make it as easy as possible for them to do so.

Unfortunately, the Issue Management function of private GitHub repositories–like the EKGF/ekg-mm repository–can only be used by people who have a GitHub account (see 3.1).

As someone in the role of a critic, you don’t have to deal with ”Agile” or ”Kanban”. The various workgroups/teams will plan for (see 4) and process/fix (see 5) your issues. You can always see the status of your issue by clicking on it and checking the right side of the screen where you see links to the ”Assignee” (the person who is going to work on your issue), the ”Project” (the workgroup that’s planning and tracking your issue) and the ”Linked pull requests”.

3.1 How to register?

We are asking every potential contributor, even someone who just wants to give us some feedback (a ”critic”) to join the EKGF.

The EKGF supports free membership for individuals and corporate membership.

Email with registration@ekgf.org to set it up, if you plan to participate as a contributor to the EKG/MM then please also supply your GitHub user-id.

3.1.1 How to create a GitHub & Slack account?

Slack If you already have a Slack account then join the EKGF workspace here: `https://ekgf.slack.com/`. Otherwise first create your Slack account here: `https://slack.com/get-started#/create`.

GitHub Go to `https://github.com/join`.

Since you can associate multiple email addresses to your GitHub account we would suggest to initially create it with your private email address which then becomes your ”primary email account” in GitHub. You can add your business email address(es) to your GitHub account later.

Mail your GitHub account id (user-id) to registration@ekgf.org so that we can add you to the access control list of the EKGF/ekg-mm repository.

3.2 How to read the content?

If you're interested in the latest version of the EKG/MM document, which is published as a PDF document, then go to `https://www.ekgf.org/maturitymodel` and download it from there.

NOTE: At the moment we do not yet put the latest version of the generated PDF document on that webpage automatically so it can take a few days until EKGF staff manually uploads the latest version to that webpage. We're working on automating that last step of the "Continuous Integration (CI) Process".

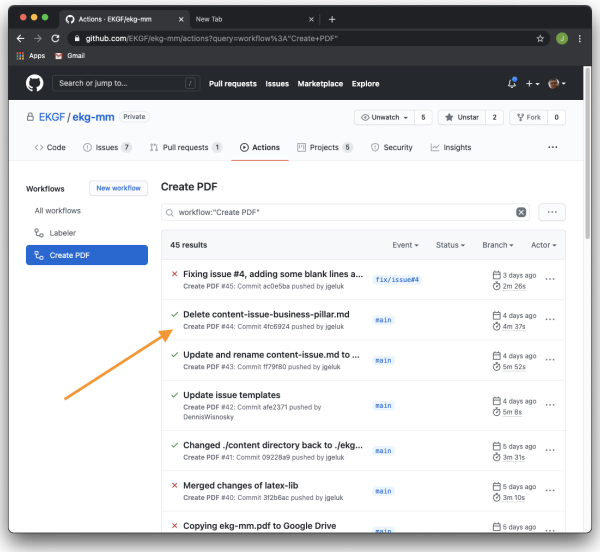


Figure I.2: Find your PDF

3.2.1 How to read specific versions of the PDF?

For readers who would like to see "work in progress", versions of the document that are being worked on in the various "branches" of the repository, there's another way to get to the corresponding PDF document.

To see the latest content for any given branch, go to the "Actions" page on the GitHub website (`https://github.com/ekgf/ekg-mm/actions`) where you can find the CI-build processes that create the PDF for each change (in git jargon that would be called "a push") in any branch.

In figure I.2 you can see how you can find the latest successful run of the "Create PDF" workflow on the `main` branch (job #44 in this case, the branch name is shown in blue).

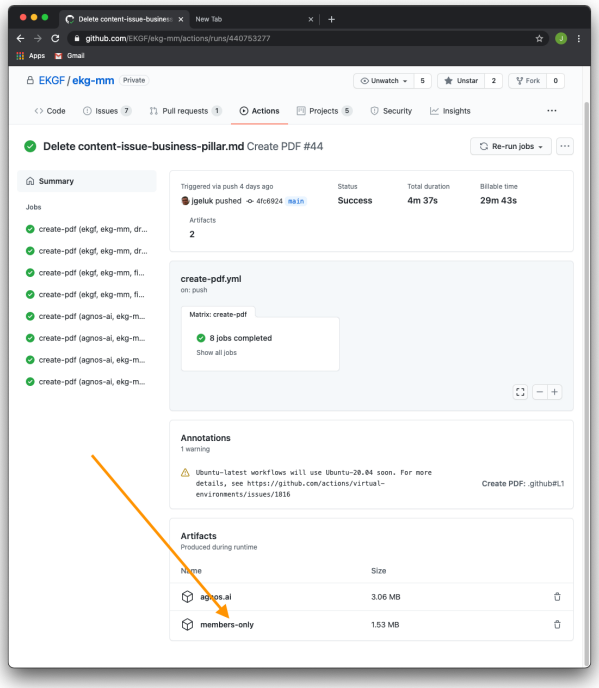


Figure I.3: The "Create PDF Job" page where you can find your PDF

When you then click on the name of that job, you will see the details of that job. When the job ran successfully, "everything is green" as you can see in figure I.3. Scroll to the bottom where you will find the "members-only" link which gives you a downloadable zip file with the "editors-version" and "release-version" of the document.

3.2.2 Document File Names

The "create-pdf" job as shown in figures I.2 and I.3 generates two different versions of the PDF:

1. `ekgf-ekg-mm-<version-number>.pdf`
2. `ekgf-ekg-mm-editors-version-<version-number>.pdf`

The documents with "`editors-version`" in the file name have a wider right margin where all kinds of "notes" and "todo's" can be shown (that are part of the content in git). They can also contain new sections of the document that are not fully consistent yet.

3.3 How to raise an issue?

While you're reading the document, you may have comments, we hope you do. Please raise as many "issues" as possible, any feedback is welcome. An issue can be a question, a suggestion, a grammar issue, an issue with the consistency of the story, anything that comes to mind. If you decide to let us benefit from your feedback then please follow these steps:

- Go to the main page where all issues are shown: `https://github.com/ekgf/ekg-mm/issues`
- Check if your issue is already raised by someone else, if so, feel free to augment the discussion around that issue with your own concerns.
- If it clearly is a new issue, click the "New Issue" button.

- Give it a concise title and clear description and ideally:
  - Copy paragraphs you don't agree with into the issue
  - Mention the paragraph number and version of the document (as shown on the front page)
- There is no need to specify values for any of the fields on the right side because the admin team will do that (but feel free to).
- Save by clicking the "Submit new issue" button

That's it. Thank you very much. Every issue is a discussion page, other people can respond to it and chip in. At some point, a workgroup will pick it up, assign it to their project board, plan for fixing it, etc. You will be notified of these changes automatically via GitHub email.

4 Plan

User Stories

- As a planner, I want to be able to **plan issues**

Work on the EKG/MM is divided among workgroups that each take on an MM pillar or horizontal slice of the MM. Workgroups are self-managed teams. The workgroups meet on a regular schedule. Any method that the team agrees to use to plan and to document issues during working sessions is ok. But, the official content of the MM is contained in the GitHub repository that is managed using the GitHub Kanban with review and automation process.

Here are some useful links.

- Kanban- A brief introduction
- What is a Kanban Board?

The idea is very simple though, we use the GitHub Projects facility. Currently, we have a project for each pillar and a general project. Here's the main page for EKG/MM GitHub Projects:

`https://github.com/ekgf/ekg-mm/projects`

What you see is a list of "project boards" where each project board is owned by a workgroup in the entire Kanban process. This begins with planning.

The planning process consists of various tasks:

- Find open issues that are relevant to the workgroup in the list of issues:
  - Here is a link: open issues that have not yet been assigned to a project board
- Link each issue that should be done by the workgroup to its corresponding project board. It will then appear at the bottom of the "To Do" column of that board.
- Decide priority (see 4.1).
- Assign issues to members of the workgroup.
  - By linking the issue to one or more assignees (open the issue by clicking on it and select the assignee at the top right corner)
- Follow the progress of issues as they go from left to right via the columns on the board.

4.1 "To Do" Column

As soon as an issue has been assigned to a given project, it will show up at the bottom of the To-Do Column. That To-Do Column can be seen as "the backlog" for that project (and for the workgroup that owns that project).

Issues can be dragged and dropped within the To-Do Column. Issues at the top have the highest priority. Anyone in the workgroup can pick up issues assigned to them and start "fixing" them (which means resolving them), see section *5, "Fix"*.

4.2 ”In progress” Column

Pull Requests and Issues[1] show up in this column automatically as soon as someone is actually working on the issue, see section *5, “Fix”*.

This means that as soon as a fixer ”pushes” one or more ”commits” to the GitHub repository that this will be seen as progress and show up accordingly on the project board.

4.3 ”Review in progress” Column

Once the fixer has created their Pull Request (PR), they can ask for a review of the PR by selecting two or more reviewers. See section *6, “Approve”* for information about how to do a review.

4.4 ”Reviewer approved” Column

Once all reviewers have approved the PR, the last reviewer can then merge the PR into the main branch. By the way, a reviewer can also reject a change or ask for changes to be applied before approving the PR.

The fixer then has to then apply new changes and commit them to the same branch and push those changes to GitHub. The PR page will then be updated automatically and a subsequent review process can then commence. As said above, once all approvals are given, the changes can then be merged by the final reviewer into the main branch which will trigger the final build workflow on GitHub creating the publishable content as PDF (and eventually also as HTML).

4.5 ”Done” Column

When the merge of the PR is done the issue will move to the Done column.

For each PR there is an underlying git branch which will be automatically deleted by GitHub once the PR has been merged into the main branch.

[1] Issues are sometimes also called “Cards” in Kanban terminology and in the Github project user interface. Besides that you can also add “Notes” to the project board.

5 Fix

User Stories

- As a fixer, I want to be able to **fix issues**
- As a fixer, I want to be able to **create a pull request**

The term ”fixer” is our own term, they’re usually called ”developer” but we don’t consider someone who works on EKG/MM content to be a developer necessarily. In the GitHub user interface, they show up as ”contributors”. Anyone who actually changes things in the repository will be registered by GitHub as a contributor to that repository[2].

There are many different ways to change the content of the repository. There are roughly two methods though:

(A) Change content without making a ”clone” of the repository by editing content on the GitHub site itself.

(B) Make a so-called ”git clone” of the repository to your own workstation and use any appropriate editor to add or change things.

Both methods require:

- ”creating a branch” (See 5.2)
- ”committing a change” (See 5.3)
- ”creating a pull request” (See 5.4)

We will now describe how each of these three tasks can be done.

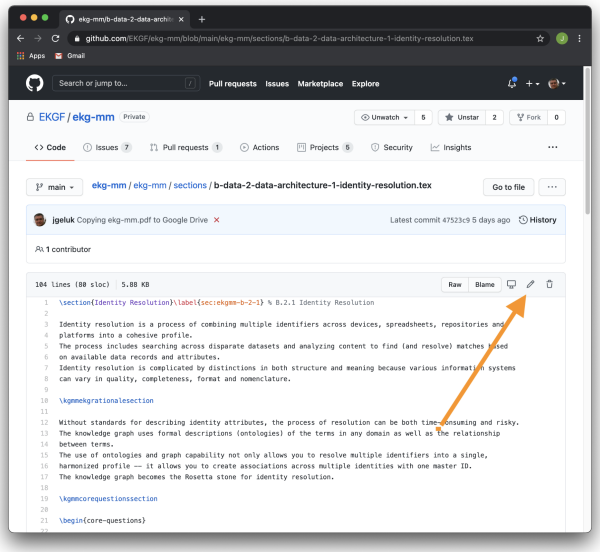


Figure I.4: Where to find the edit button

5.1 How to start editing in GitHub?

This is ”Method A” as described above. First find the content file that you want to edit. All content is under the content root directory `/ekg-mm`.

The arrow in figure I.4 above shows the location of the ”edit button” (ignore the other buttons). Click that button and change any content. All content is written in a ”language” called LaTeX which is a markup language for professional content (used for books and Ph.D. thesis and the like). As long as you keep the LaTeX commands (aka macros) in place (they start with a backslash) things will be fine, just type away. Unless you know LaTeX (a little) then you can add whatever works.

Then it’s time to save the changes, choose a good title for your change, and include the issue number (for instance #123) prefixed with a hash. This title is the so-called ”commit message” and is described in more detail in section *5.3, “How to commit a change?”*.

5.2 How to make a branch?

When editing straight on the GitHub site as described in the previous section, you have to save your work ”as a commit” with a title that contains the issue number (which will be picked up by the project board). But you then also have to create a branch which shows up right under the title and description.

Choose a branch name that explains what it is about. In some cases that would be all there is to it but in most cases you have multiple edits in multiple files for one given issue. So in those cases you would have to make sure to always refer to the right branch name where the branch name must include the issue number itself. For instance, for issue #123 the branch name would be `issue-123`.

5.3 How to commit a change?

[2] The GitHub userids and names of all contributors will be listed in the final output EKG/MM document.

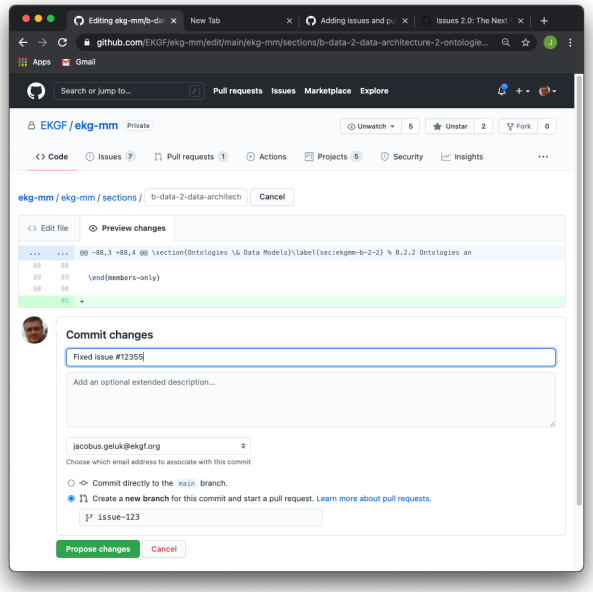


Figure I.5: Committing changes

There are many different ways to create a commit on a branch in a git repository. It's up to the "fixer" (usually called "developer" in most git documentation on the Internet) to decide which tooling they use. One way to do this is to just use the Github website itself but you could also create a so-called "clone" of the git repository on your local machine and use the standard git utility to create a commit or use an advanced editor like Visual Studio Code that has built-in git support and a LaTeX plugin. It goes too far to document all the various different ways to do this in the context of this document. What's important to know is that, as a fixer, you should always include the issue number in the title of the commit. For example, if you edit a content page and then commit that change, you have to make sure that these two things are done before clicking the "Propose Changes" button as shown in figure I.5.

Use a title for your commit that includes the issue number, prefixed by #, for example, #123 for issue 123. Use a branch name that includes the issue number (no hash as a prefix).

The #123 in the title will trigger GitHub to move the corresponding issue from the "To Do" column to the "In Progress" column. Do not use "Fixed #123" or "Close #123" because that will actually close the issue.

Once you have done your first commit to a given branch, you're ready to create a PR. Many people wait with creating a PR until their last commit is pushed (one issue can involve many commits involving many files but they all go to the same branch) but we recommend creating your PR as soon as possible, after the first commit, because that gives everyone an easier insight in the progress. Not everyone is up for that, "exposing" their first draft changes to others, but it's generally seen as good practice to do so. You can mark your PR as "Draft" via the GitHub user interface so that everyone knows that it is "work in progress".

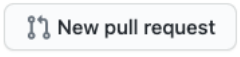
5.4 How to create a pull request?

Assuming that you added some changes to a branch, in GitHub parlance: "pushed some commits to a branch", you can also create a PR. A PR can be seen as a "request approval to change"-form. It needs two branches: the branch that you want to change (that would usually be the main branch) and the branch that contains the new version with the proposed changes. GitHub automatically calculates the difference between those two branches and shows those differences on the PR-form.

There are several different places in the GitHub user interface where you can start creating a PR but the most basic one shows up at the branches page which shows all branches of the repository:

`https://github.com/ekgf/ekg-mm/branches`

On the right side of the table showing all the branches you either see an existing PR for the given branch or a button:



It's easy from there. Give your PR a good title (that title goes into the "changelog" so be specific) and a description and click the Create Pull Request button. Job done. A planner in the workgroup will take it from there, assign reviewers, link it to a project board etc.

6 Approve

User Stories

- As a reviewer, I want to be able to **approve pull requests**

TODO

Appendix J

History

Version History

Version	Date	Description	Author
0.1	01/03/2019	agnos.ai started with the EKG/MM initiative	Michael Atkin, Jacobus Geluk
0.1	01/01/2020	agnos.ai donated IP to EKG/MM to EKGF “as is”	Jacobus Geluk
0.2	01/01/2021	Started with new delivery process	Mike Atkin, Dennis Wisnosky
0.3	08/04/2021	Business pillar workgroup deciding top-level structure	Lars Magnusson, Carl Mattocks, Jeffrey Wallk *et al*.
0.4	27/05/2021	New intro for business pillar	Avinash Patil, Carl Mattocks, *et al*
		Content for A.2.3 Change Management	Diane Alexander, Jeffrey Wallk
		Removed chapter ”A.1.4 Operating Model”	
		Added content for chapter ”A.3.3 Risk Management”	Diane Alexander
0.5	08/06/2021	Added appendix ”What is Strategy?”	Carlos Tubbax
		Added several updates to Business Pillar intro	Avinash Patil
0.6	14/06/2021	Added chapter ”A.3.1 Operating Model”	Carlos Tubbax
0.7	01/07/2021	Moved operating model version of Carlos to Appendix and added a new appendix ”Enterprise Architecture”	Carlos Tubbax
0.8	31/08/2021	Added completely new content for ”B.2.2 Ontologies”	Pete Rivett and others
		Added completely new content for ”A.3.2 Enterprise Architecture”	Carlos Tubbax
		A.3.2 Performance Management is now A.3.3 Performance Management	
		A.3.3 Risk Management is now A.3.4 Risk Management	
		A.3.4 Supply Chain Management is now A.3.5 Supply Chain Management	
		Removed A.3.2 Enterprise Architecture, moved it to appendix	
0.9	05/10/2021	Reshuffle of various capabilities in Data Pillar	Pete Rivett
0.10	13/10/2021	Moved Organization Pillar to second place	Avinash Patil
0.11	26/10/2021	Added chapter A.3.5 Capability Map	Carlos Tubbax

Appendix K

Contributors

All contributors in alphabetical order.

Contributor	Org	Role
Diane Alexander	Factor	Business Pillar
Michael Atkin	Columbia University	EKGF Founder EKG/MM Initiator Data Pillar
Marcel Fröhlich	eccenca GmbH	EKGF Board All pillars
Jacobus Geluk	agnos.ai	EKGF Founder EKG/MM Initiator All pillars
Howard Knowles	agnos.ai	Technology Pillar
Ross Leher	WAND	Business Pillar
Mary Levins	Sierra Creek Consulting	Business & Data Pillar
Lars Magnusson	Ericsson	Business Pillar
Carl Mattocks	GE, New York Power Authority	Business Pillar
Eduardo Martinez	BMO	Business Pillar
Avinash Patil	Tata Consultancy Services	Business Pillar
Pete Rivett	OMG agnos.ai	EKGF Founder & Vice President Data Pillar
Marina Severinovskaya	Citi	Business & Data Pillar
Carlos Tubbax	University of Antwerp	Business Pillar
Jeffrey Wallk	Value Enablement Group	Business Pillar
Dennis Wisnosky	Wizdom	EKGF Founder & Chair EKG/MM Process
Jeannine Wisnosky Stehlin	Wizdom	EKGF Director of operations EKG/MM Process
Omar Bryan Khan	Wells Fargo	Technology Pillar

Appendix L

Corporate Members

The corporate members of the Enterprise Knowledge Graph Foundation (EKGF) currently are — ordered by join date:

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

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

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

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

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

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

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

Cambridge Semantics Cambridge Semantics provides Anzo, the scalable knowledge graph platform for data integration and analytics.

Acronyms

AI Artificial Intelligence 10, 25, 56
API Application Programming Interface 47, 49, 68
ASCM Association for Supply Chain Management 30
BAU Business-as-Usual 9, 10, 19, 21
CDE Critical Data Element 53, 59
CI Continuous Integration 99
CIO Chief Information Officer 13, 35
COSO Committee of Sponsoring Organizations of the Treadway Commission 29
EKG Enterprise Knowledge Graph 1, 6–11, 13, 16–25, 27, 29, 30, 32, 37, 47–59, 67, 68, 73, 74, 76–83, 94, 96, 106, 108, Glossary: EKG
EKG/Manifesto Manifesto & Principles for the Enterprise Knowledge Graph 81, 94
EKG/Method Enterprise Knowledge Graph Method 81, 94
EKG/Platform Enterprise Knowledge Graph Platform General Terms: EKG/-Platform
EKG/MM Maturity Model for the Enterprise Knowledge Graph 6–8, 81, 96, 97, 99, 100, 102
EKGF Enterprise Knowledge Graph Foundation 6, 77, 80, 81, 96–99, 105, 106, 108
EPM Enterprise Performance Management 28
ERM Enterprise Risk Management 29, 30
ERP Enterprise Resource Planning 90
ESG Environmental, Social, and Corporate Governance 15
ETL Extract Transform Load 53, 69
FIGI Financial Instrument Global Identifier 77
IaC Infrastructure as Code 10
IETF Internet Engineering Task Force 108
IoT Internet of Things 25
KPI Key Performance Indicator 50
LEI Legal Entity Identifier 77
LOB Line of Business 9, 31, 32
MCDA multiple-criteria decision analysis 25
MVOT multiple versions of the truth 73, 77
NFR non-functional requirement 81
NLP Natural Language Processing 51
ODM Office of Data Management 57
OWL2 Web Ontology Language 77
PDCA Plan, Do, Check, Act 16, 23
PoC Proof of Concept 9
PR Pull Request 101, 103
RDF Resource Description Framework 76, 77, 108
ROI Return On Investment 13, 32, 45
S3 Simple Storage Service Glossary: object store
SDD self-describing dataset 54, 74, 108, Glossary: self-describing dataset (SDD)
SLA Service-level Agreement 52
SMART Specific, Measurable, Achievable, Realistic, Timely 19, 21
SME Subject Matter Expert 50, 56
SOA service-oriented architecture 32, 68
SOR System of Record 52, 58, 59
SUDA Sense, Understand, Decide, Act 16, 19, 23
SVOT Single Version of (the) Truth 47, 73
UCT Use Case Tree 32, 52, 81
UNA Unique Name Assumption 77
W3C World Wide Web Consortium 77, 108

Glossary

CoE for the EKG the group of people that is overseeing the EKG 9, 10, 49, 57, 81, 82, 94, 108

EKG use case TODO 49, 50, 79, 82

Canonical Identifier A permanent identifier for an object, distinguished within an EKG. 108

data point a fact representing the value(s) of a property for an object in some context. Each value may be a simple data (e.g. a number, string or date) or another object. The combination of object, property and value is a *triple*. 74, 77–79

dataset a collection of data, published or curated by a single agent, and available for access or download in one or more serializations or formats. 78, 79

EKG much like "the web" is a virtual concept, an EKG is a virtual concept that combines all information and knowledge of an enterprise — at any level in the organization — or ecosystem. 6, 67, 73, 76,

EKG system architecture the logical system architecture of EKG is divided into multiple layers or environments such as EKG/Platform, EKG/Storage and EKG/DataOps environment. 108

EKG/DataOps environment a logical system architecture component, the environment where EKG/DataOps pipelines run. Since EKG/DataOps pipelines can run anywhere and any given data provider can publish their own data as a SDD in any technical or physical environment, the actual DataOps "environment" is just a logical concept that does not necessarily map one-to-one to a particular physical environment. 108

EKG/DataOps pipeline a DataOps pipeline in the EKG/DataOps environment is a series of programs, called "steps", that are run in sequence — hence the name pipeline — where the first step captures the data from a given source and the last step produces an SDD (for "inbound pipelines") or any other type of artifact (for "outbound pipelines"). Each DataOps pipeline step takes an input and produces an output based on its configuration which usually refers to a model that instructs the DataOps pipeline step how to process the data. Crafting these models is the main activity in the EKG/DataOps Practice. 108

EKG/IRI an IRI that forms the identity of an object in the EKG. Any given object in an EKG has an EKG/IRI for which special rules are defined by the EKGF. Not to be confused by Canonical Identifiers. 76, 77, 108

EKG/Method a methodological approach to the development of an Enterprise Knowledge Graph (EKG), covering all practices from the definition of desired business outcomes to deployment of EKG use cases that deliver on these outcomes and beyond. 108

EKG/Platform a logical system architecture component (see EKG system architecture), the layer of software services that provide and serve the EKG to end-users and other systems. The platform logically is a set services that enforce any of the specified policies in the SDDs that have been published in the EKG. 9, 10, 13, 39, 40, 64, 65, 67–69, 74, 77, 107, 108

EKG/Storage a logical system architecture component, the layer of data-storage services such as a "Triplestore" or an "Object Store" that serve the various other layers in the EKG system architecture. 108

IRI a standard defined by the Internet Engineering Task Force (IETF) in 2005 in RFC 3987. An Internationalized Resource Identifier (IRI) is defined similarly to a , but the character set is extended to the Universal Coded Character Set. Therefore, it can contain any Latin and non-Latin characters except the reserved characters. Instead of extending the definition of , the term IRI was introduced to allow for a clear distinction and avoid incompatibilities. IRIs are meant to replace URIs in identifying resources in situations where the Universal Coded Character Set is supported. By definition, every URI is an IRI. Furthermore, there is a defined surjective mapping of IRIs to URIs: Every IRI can be mapped to exactly one URI, but different IRIs might map to the same URI. Therefore, the conversion back from a URI to an IRI may not produce the original IRI. The IRI standard is a superset of the older "Uniform Resource Identifier (URI)" standard (IRI ⍰ URI). See also EKG/IRI. 76, 77, 108

lighthouse project Lighthouse projects simply focus on implementation, fast delivery and creating a positive culture for digital transformation. A lighthouse project is a short-term, well defined, measurable project that serves as a model — or a "lighthouse" — for other similar projects within the broader digital transformation initiative. See Why Lighthouse Projects, Not PowerPoints, Will Unlock Your Transformation Value 9

linked data the collection of interrelated datasets on the Web. Linked Data is a concept defined by Tim Berners-Lee and the W3C that is part of what they call "the Semantic Web" or "the Web of Data". Tim Berners-Lee started the idea in 2006 by defining 4 simple rules: `https://www.w3.org/DesignIssues/LinkedData.html`. 77

low-code is a software development approach that requires little to no coding in order to build applications and processes. In an EKG context, low-code — or actually it's more like "no-code" — development is achieved by specifying business and user requirements at such a level of detail, with or without the use of visual user interfaces for developers, that at some point smart "model executing services" can provide a running production-level use case. 82

ontologist TODO 48

operating model is both an abstract and visual representation of how an organization delivers value to its customers or beneficiaries as well as how an organization actually runs itself. 9, 10, 57, 58

practice the DataOps practice, — as defined by the EKG/Method — is one of the practices that the CoE for the EKG executes. 108

predicate asserting an "RDF triple" says that some relationship — indicated by the predicate — holds between the resources denoted by the subject and object. This statement corresponding to an RDF triple is known as an RDF statement. The predicate itself is an IRI and denotes a property, that is, a resource that can be thought of as a binary relation. 108

predicate-IRI see predicate 49

property an identified and fully-defined linkage between an object and values. Or in RDF terminology, a relation between subject resources and object resources. See `https://www.w3.org/TR/rdf-schema/#ch_property`. 77

RDF Schema TODO RDF Schema 48

self-describing dataset (SDD) an EKG is logically composed of a set of *self-describing datasets* that provide information about lineage, provenance, pedigree, maturity, quality, governance, entitlement policies, retention policies, security labels, IP policies, pricing policies, caching policies, organizational ownership, accountabilities, data quality feedback loop, issue management policies and so forth. Any given system owner that connects their data to the EKG becomes in fact a publisher of a self-describing dataset and can therefore control how their data is handled by the EKG/Platform. The EKG/Platform consists of services that enforce any of the specified policies in the self-describing dataset. At the technical level, datasets can be files or streams or API definitions etc. See also principle *5 Self-describing* on page 78. 54,

system architecture the conceptual model that defines the structure, behavior, and more views of a system. 108

technology stack ... 13, 35, 54

URI a Uniform Resource Identifier (URI) is a compact sequence of characters that identifies an abstract or physical resource. See `https://www.ietf.org/rfc/rfc3986.txt` and IRI. 108

References

[2021-Risk] Enterprise Risk Management Initiative Staff. *2021 The State of Risk Oversight: An Overview of Enterprise Risk Management Practices*. 12th Edition. 2021. URL: https://erm.ncsu.edu/library/article/2021-risk-oversight-report-erm-ncstate-lp (cit. on p. 29).

[buscap] William Ulrich and Michael Rosen. *The Business Capability Map: The "Rosetta Stone" of Business/IT Alignment*. Link to PDF. 2011. URL: https://www.cutter.com/article/business-capability-map-rosetta-stone-businessit-alignment-469506 (cit. on p. 31).

[EKG/MM] EKGF and agnos.ai. *Maturity Model for the Enterprise Knowledge Graph*. The Enterprise Knowledge Graph Foundation® (EKGF®) is an international, open membership, not-for-profit EKG standards organization. 2020. URL: https://www.ekgf.org/maturitymodel (cit. on p. 8).

[EKG/P] EKGF. *EKG/Principles*. Foundational principles for the development and deployment of an Enterprise Knowledge Graph. The Enterprise Knowledge Graph Foundation® (EKGF®) is an international, open membership, not-for-profit EKG standards organization. 2019. URL: https://www.ekgf.org/principles (cit. on pp. 9–11).

[ENISA] *Integration of Risk Management with Business Processes*. 2009. URL: https://www.enisa.europa.eu/topics/threat-risk-management/risk-management/current-risk/business-process-integration (cit. on p. 29).

[ERM-COSO] *Enterprise Risk Management Integrating with Strategy and Performance*. Executive Summary. 2017. URL: https://www.coso.org/Documents/2017-COSO-ERM-Integrating-with-Strategy-and-Performance-Executive-Summary.pdf (cit. on p. 29).

[Hirt] Martin Hirt. *If you're not building an ecosystem, chances are your competitors are*. URL: https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/the-strategy-and-corporate-finance-blog/if-youre-not-building-an-ecosystem-chances-are-your-competitors-are# (cit. on p. 15).

[Kalim] Faisal Kalim. *The Subscription Economy has grown over 435% in 9 years (and the uptick is expected to continue)*. URL: https://whatsnewinpublishing.com/the-subscription-economy-has-grown-over-435-in-9-years-and-the-uptick-is-expected-to-continue/ (visited on 03/05/2021) (cit. on p. 15).

[Porter] Michael E. Porter. *What Is Strategy?* Harvard Business Review. URL: https://store.hbr.org/product/hbr-s-10-must-reads-on-strategy-including-featured-article-what-is-strategy-by-michael-e-porter/12601 (cit. on p. 84).

[18] *The Customer Moment With Dynamic Predictive Segmentation*. a January 2018 commissioned study conducted by Forrester Consulting on behalf of simMachines. 2018. URL: https://simmachines.com/forrester/ (cit. on p. 23).

[Del20] Deloitte. *If You Want Your Digital Transformation to Succeed, Align Your Operating Model to Your Strategy*. 2020. URL: https://hbr.org/sponsored/2020/01/if-you-want-your-digital-transformation-to-succeed-align-your-operating-model-to-your-strategy (visited on 01/31/2020) (cit. on p. 26).

[Mar10] Roger L. Martin. *The Execution Trap*. 2010. URL: https://hbr.org/2010/07/the-execution-trap (visited on 07/31/2010) (cit. on p. 28).

[RWR06] Jeanne W. Ross, Peter Weill, and David C. Robertson. *Enterprise Architecture As Strategy: Creating a Foundation for Business Execution*. Harvard Business Review, 2006. ISBN: 1-591-39839-8 (cit. on p. 85).

[Tor] Simon Torrance. *Embedded Finance: a game-changing opportunity for incumbents*. URL: https://www.simon-torrance.com/blog/EmbeddedFinance1 (cit. on p. 15).

Index