midwest
architecture community
collaboration

Graph Systems Thinking

Dan McCreary

November 4th, 2021

OPTUM® | Technology

Session Description

This session will introduce the concept of Graph Systems Thinking (GST) and how it can impact your strategic thinking about the role of integration architecture in your firm. We combine developments in large-scale **Enterprise Knowledge Graphs** (EKGs) and **Systems Thinking** to show how these new ways of problem solving can give your teams a more holistic view of technology in your organization. We will show how feedback and casual loop diagrams help you understand how machine learning can empower your staff to make better prediction and better serve your customers.

Agenda

What are Enterprise Knowledge Graphs? Why are they strategic to Enterprise Architecture?

What is Systems Thinking? What are Causal Loop Diagrams and Archetypes?

What is **G**raph **S**ystems **T**hinking (GST)?

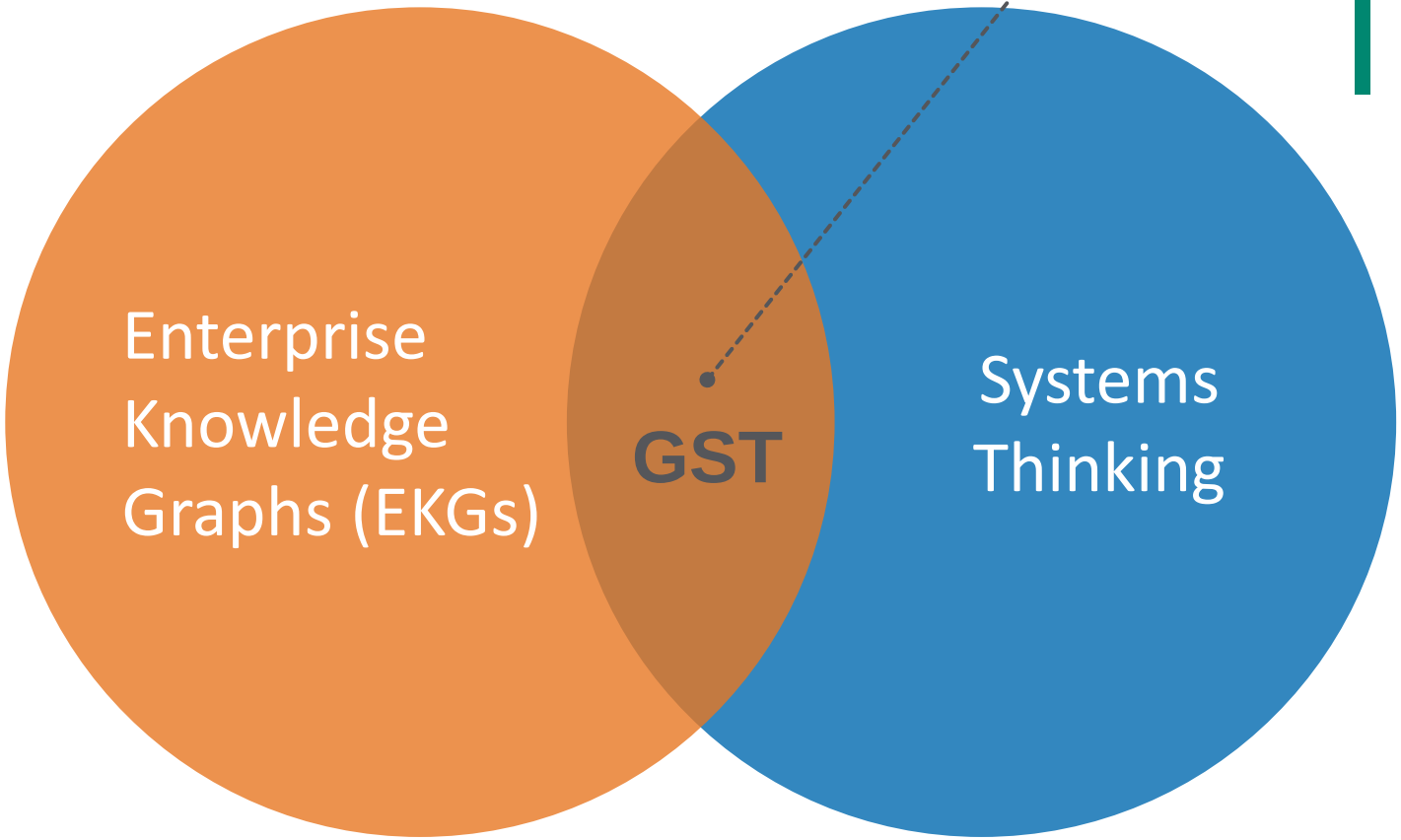
Why is GST Important for Enterprise Architecture?

GST Use Cases and Storytelling

Conclusion, Request for Help and Recommended Next Steps

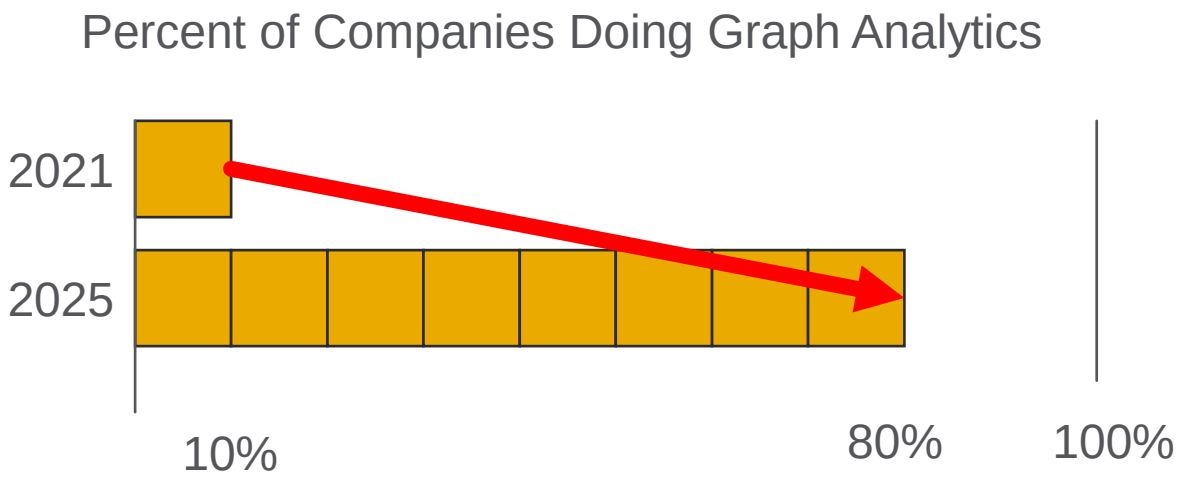
OPTUM® | Technology

EKGs and Systems Thinking



How to use **Systems Thinking** to guide the adoption of **Enterprise Knowledge Graphs (EKGs).**

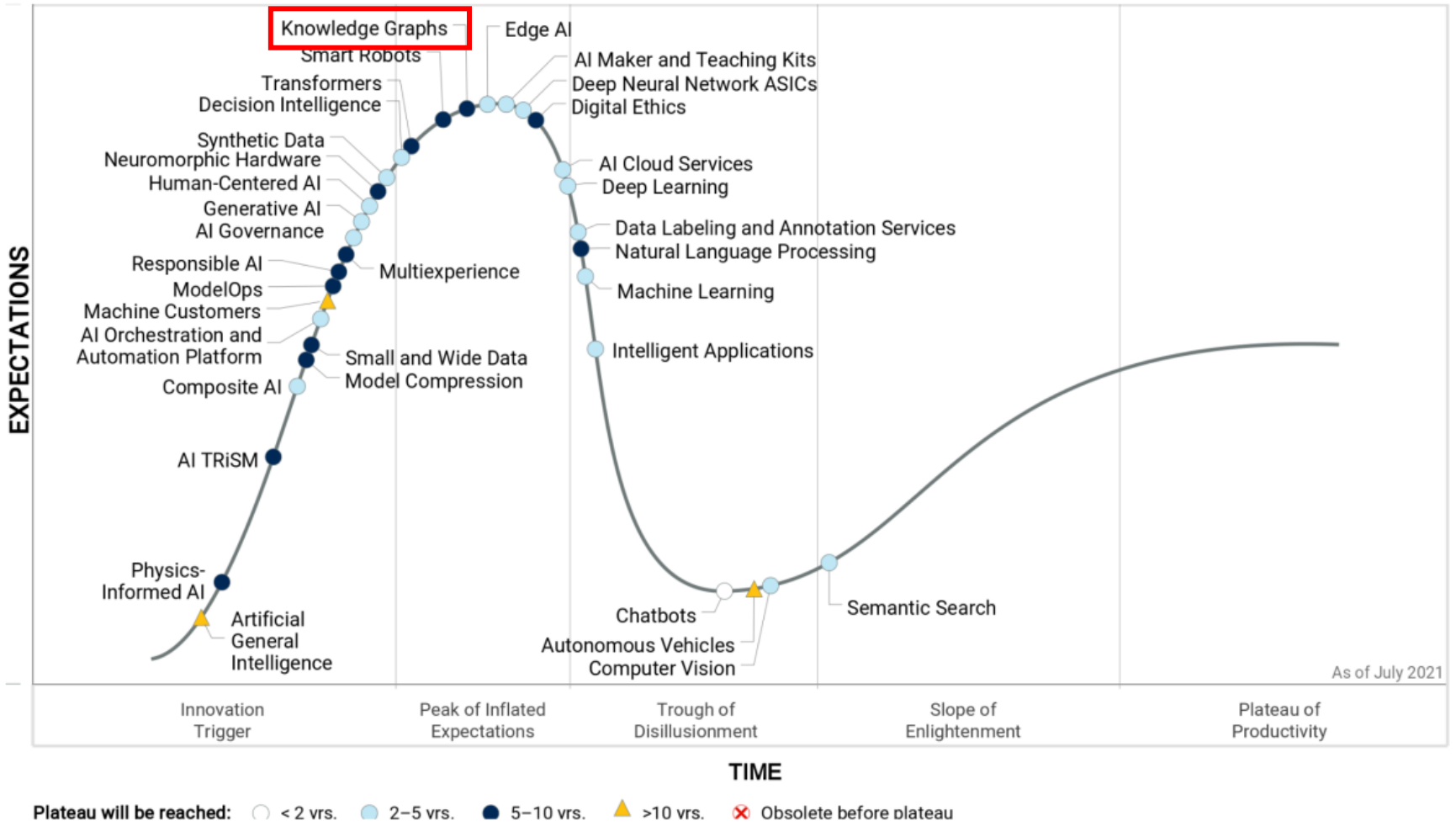
Key Trends Enterprise Architects Must Grasp



10% to 80% in just 4 years!

According to Gartner, "By 2025, graph technologies will be used in 80% of data and analytics innovations, up from 10% in 2021, facilitating rapid decision making across the enterprise.

Knowledge Graphs Are Peaking in Expectations



OPTUM® | Technology

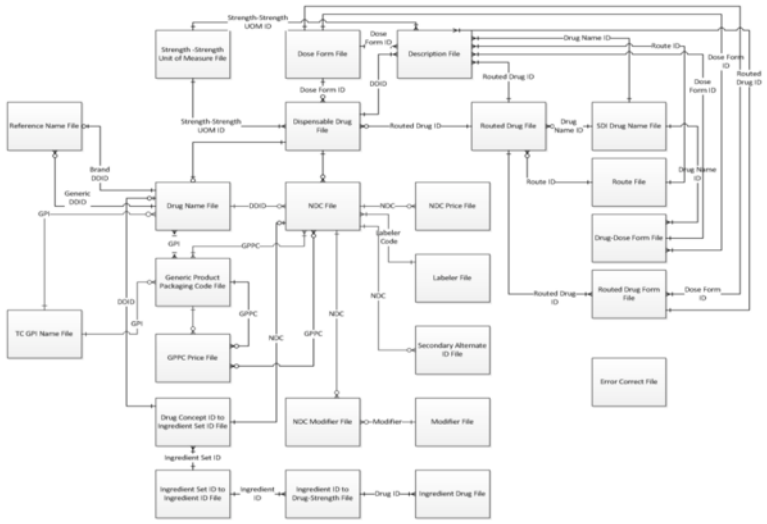
Gartner

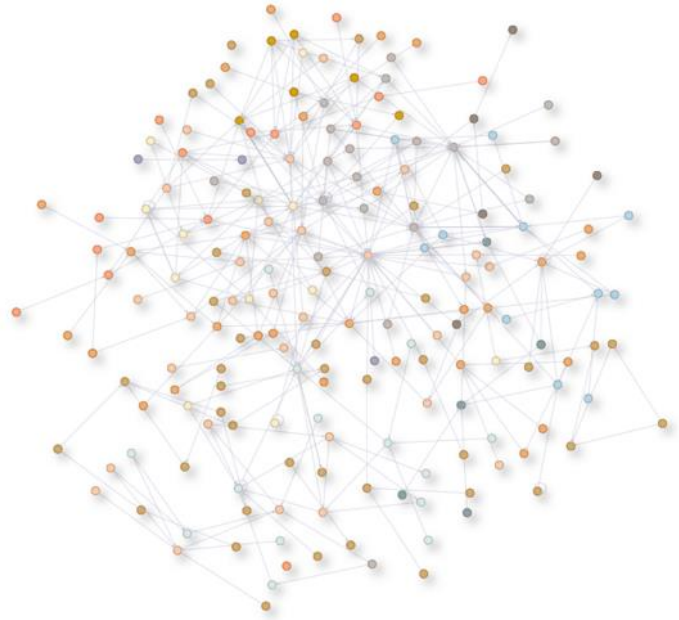
What are Enterprise Knowledge Graphs (EKGs)?

Why are they strategic to Enterprise Architecture?

OPTUM® | Technology

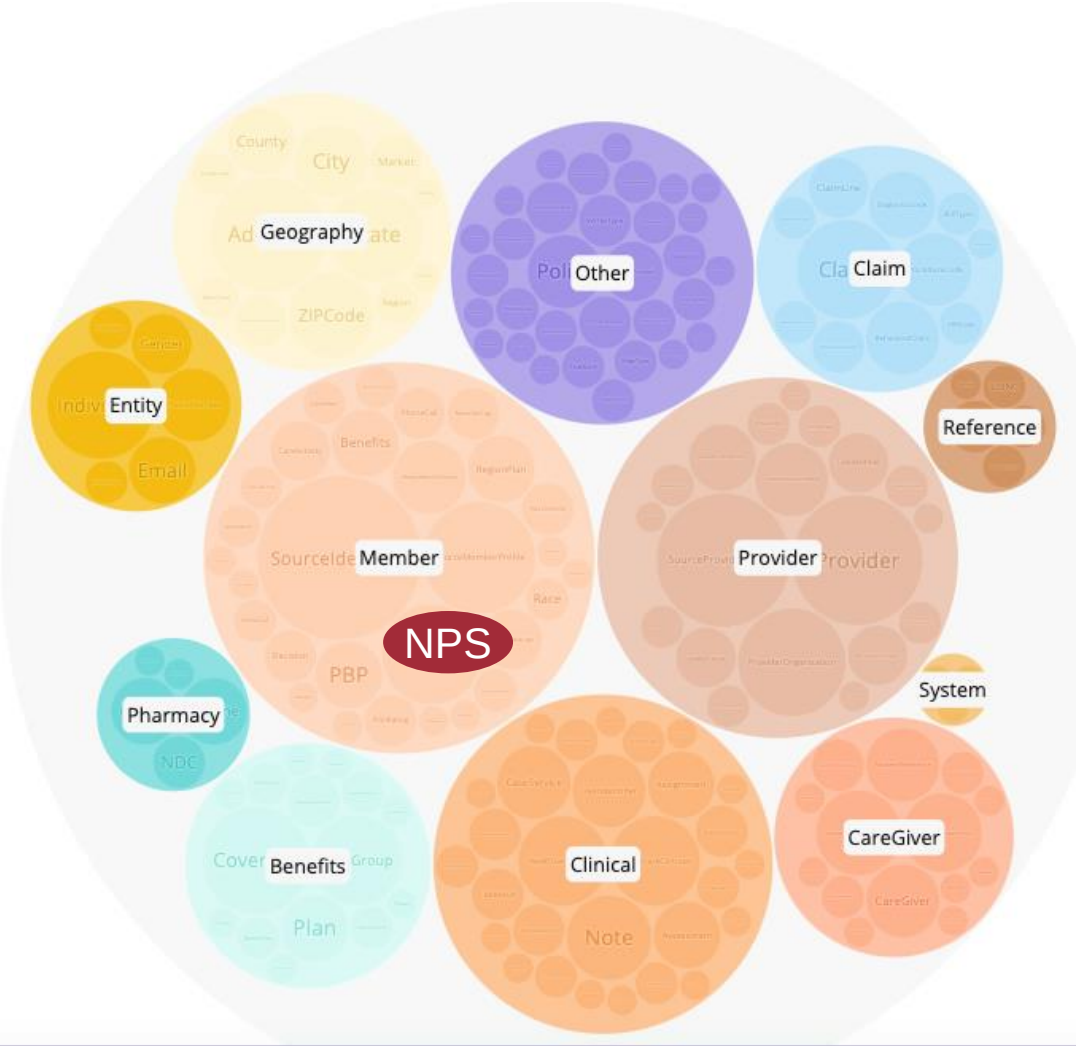
What are Enterprise Knowledge Graphs?





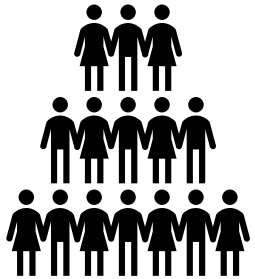
- Large 100 billion-edge graphs that span multiple business units
- Provide sub 100 millisecond response times to C360 views (10K events) to 100K concurrent users
- Designed from the ground-up to be scalable to hundreds of servers
- Designed for high-availability: easy to add more servers and disk without service interruption
- Change from modeling to minimize JOINs to model to maximize sharing

Cross Discipline Queries



- EKGs excel at bring data together from multiple business areas of a company.
- EKGs can find deep relationships between our member NPS and our internal operations.
- Example: What is the relationship between treatment T that claims Y results and the real-world results of T?

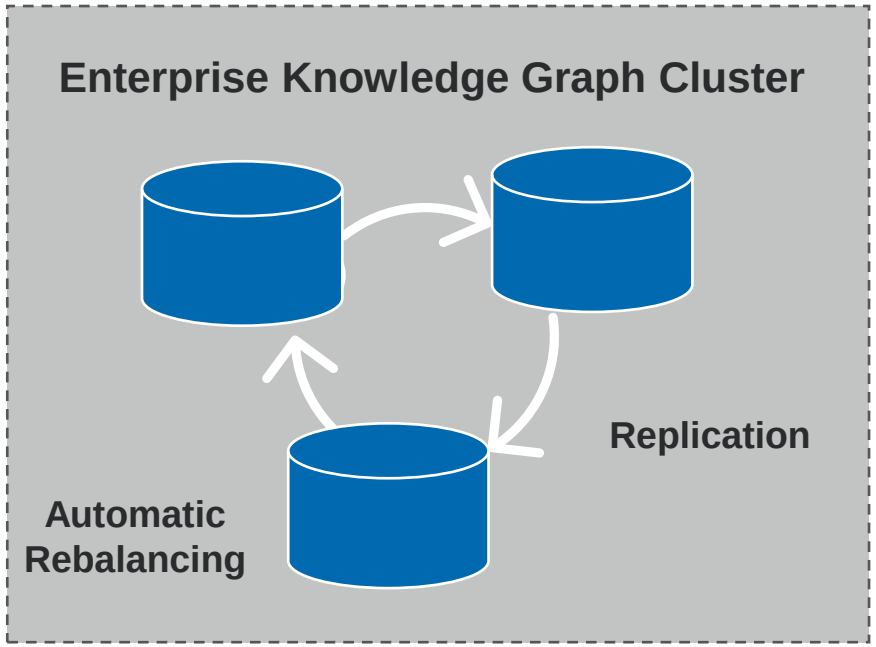
Definition of an Enterprise Knowledge Graph



Thousands of Concurrent Users

Hundreds of Applications

Secure Ad-Hoc Query



- High Availability (HA): over 99.99% uptime
- Rolling upgrades: never shut down services for upgrades
- Vertex-level Role-based Access Control (RBAC)
- Resource quotas
- Large library of graph algorithms

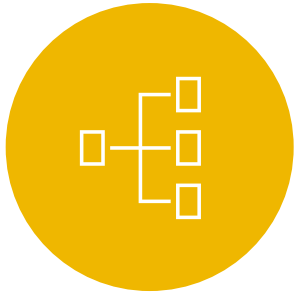
An Enterprise Knowledge Graph (EKG) is a type of graph database designed to **scale-out** to meet large organizations' demanding requirements to store diverse forms of connected knowledge.

Seven Measure of Scale-out In EKGs

Scale-out data size

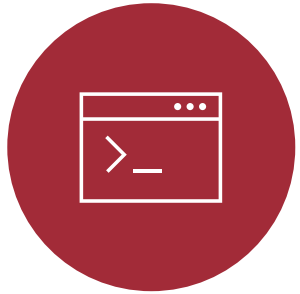
Scale-out compute

Scale-out security



Scale-out manageability

Scale-out data quality



Scale-out algorithms

Scale-out query

Beware of false prophets!

https://dmccreary.medium.com/a-definition-of-enterprise-in-ekgs-561283d37deb

EKGs: Today vs Future

Today

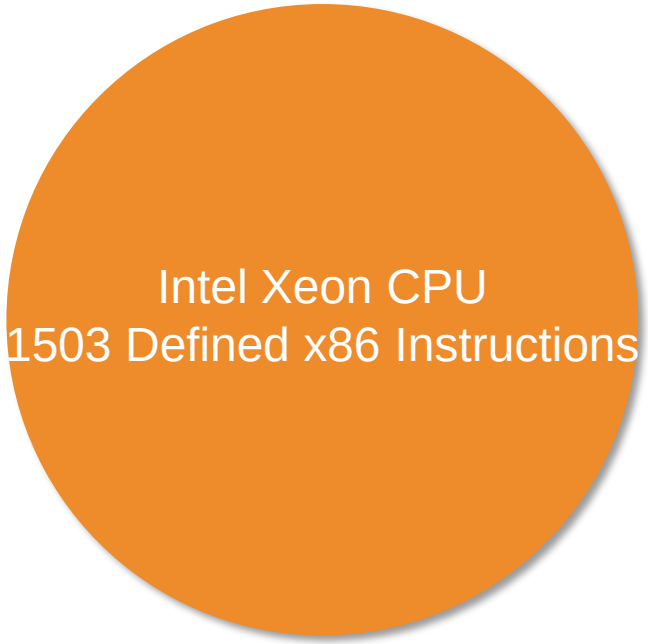
- Many graph database don’t scale well over 100s of servers
- Distributed graph database licenses are prohibitively expensive ($1M/TB/year)
- Only the largest companies can afford them
- No specialized graph chips
- No built-in machine learning

Future

- Most software vendors will have scale-out graph solutions
- Open-source distributed graphs will be common
- Even small-medium companies will have robust EKGs
- Specialized graph hardware: FPGA and Intel PIUMA
- Out-of-the box machine learning

General CPU Hardware vs. Next-Generation Graph Hardware

Complex Instruction Set Computer (CISC)

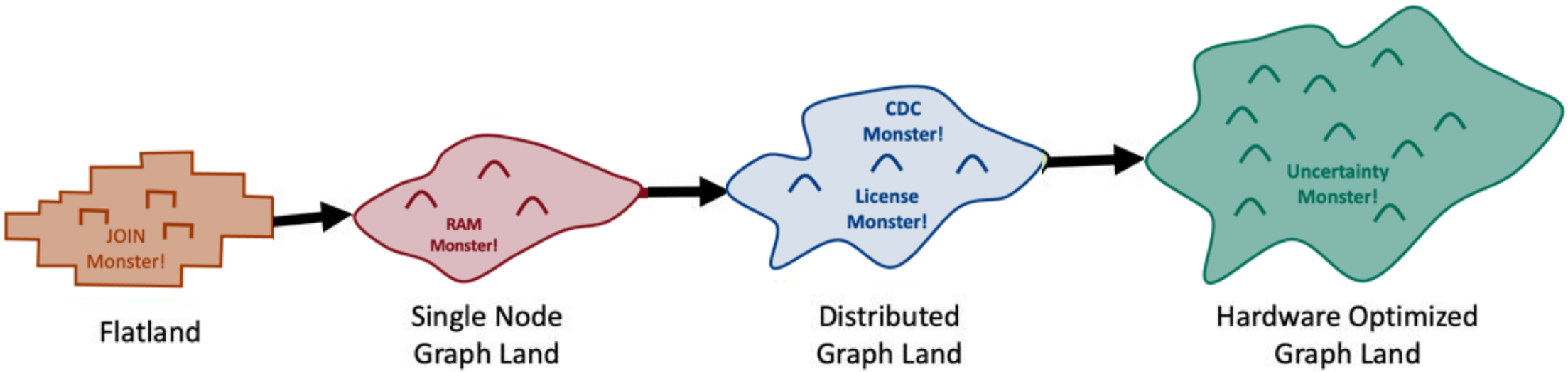


Reduced Instruction Set Computer (RISC)

Fewer Instructions -> More Cores

- Most graph traversal algorithms only need simple pointer hopping
- How efficient are CPU and GPUs at running graph algorithms?
  - No need for floating point
  - No need for matrix multiplication
- Can FPGAs be used?

Four Stages of EKG Adoption



1. Flatland
2. Single node graph
3. Distributed graph
4. HOG Heaven

[From Flatland to HOG Heaven]

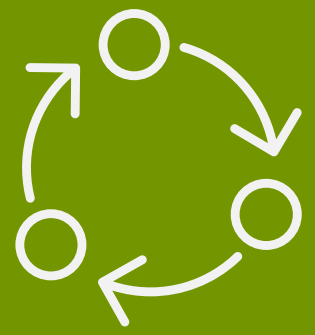
HOG Heaven

- Hardware optimized graph solutions
- 1 trillion edges are easy
- Vertex-level role-based access control
- Complete views of customers in under 100 milliseconds
- Graph embeddings for every vertex to enable fast similarity at scale

Photo by Bruno van der Kraan on Unsplash

What is Systems Thinking?

What are Causal Loop Diagrams and Archetypes?



OPTUM® | Technology

What is a System?

A collection of components that interact together to produce some sort of behavior of the whole

Systems can have subsystems

What is a “System”?

A system is a collection of components that interact together to produce some sort of behavior of the whole

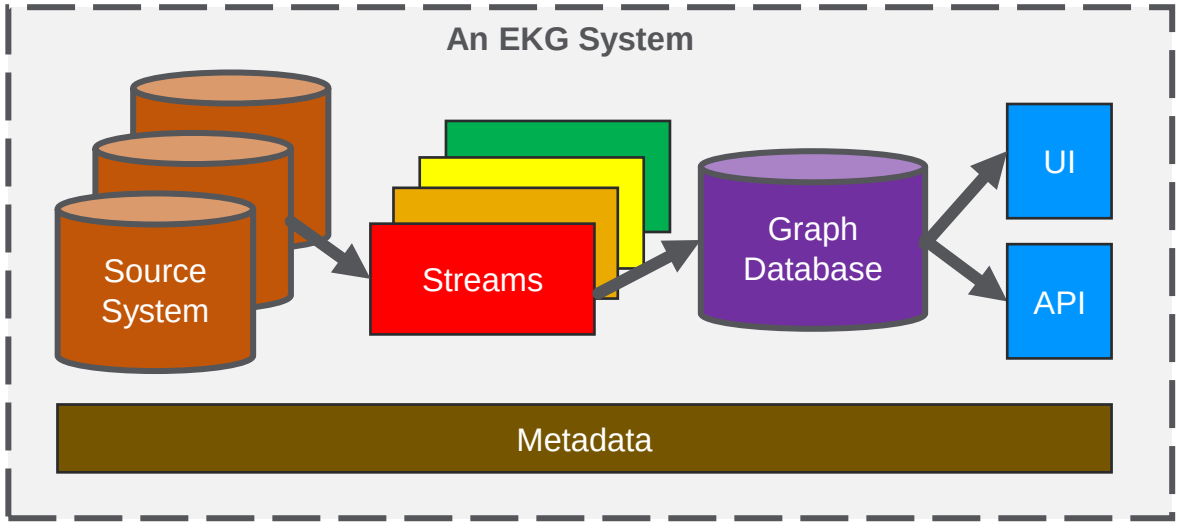
Systems can be complex

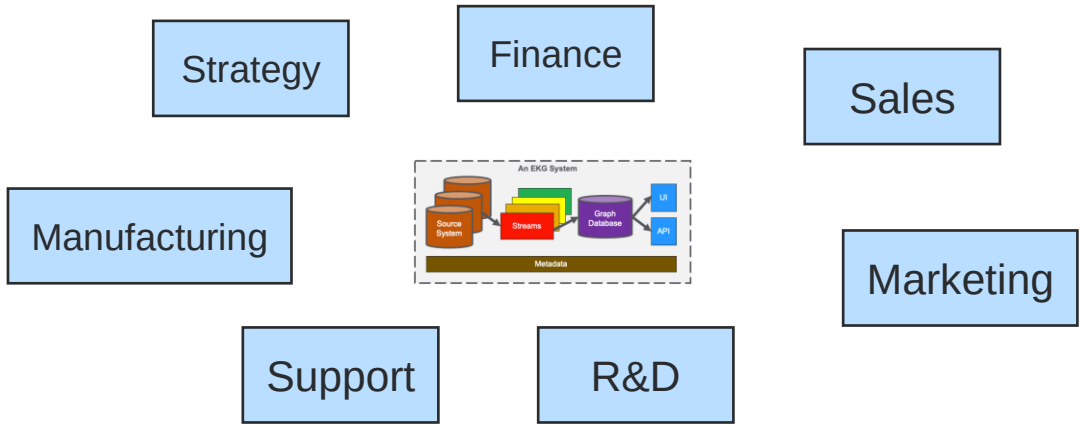
Systems can have subsystems

Systems are part of other systems

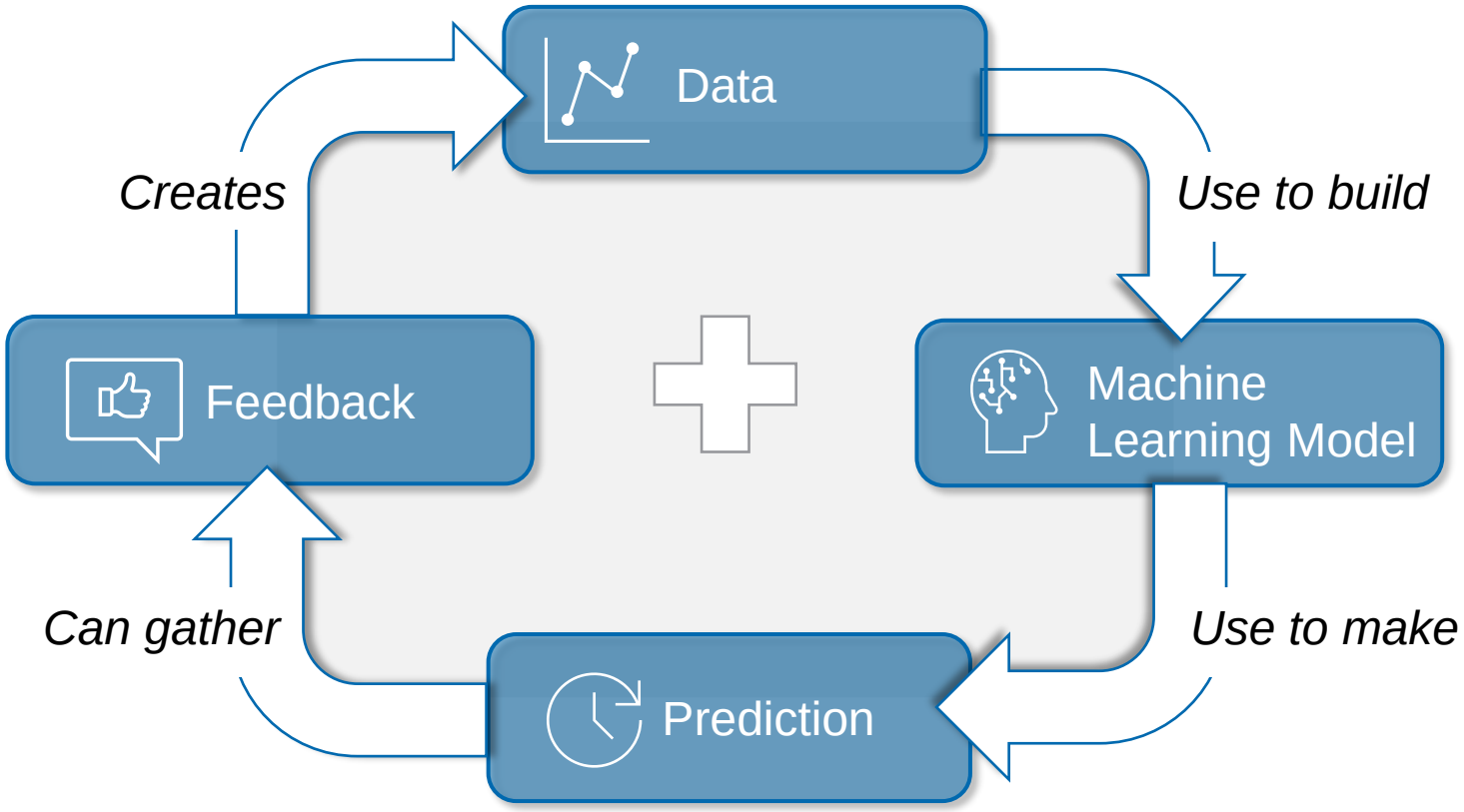
Systems evolve over time

Emergent behavior arises in complex systems





The AI Flywheel



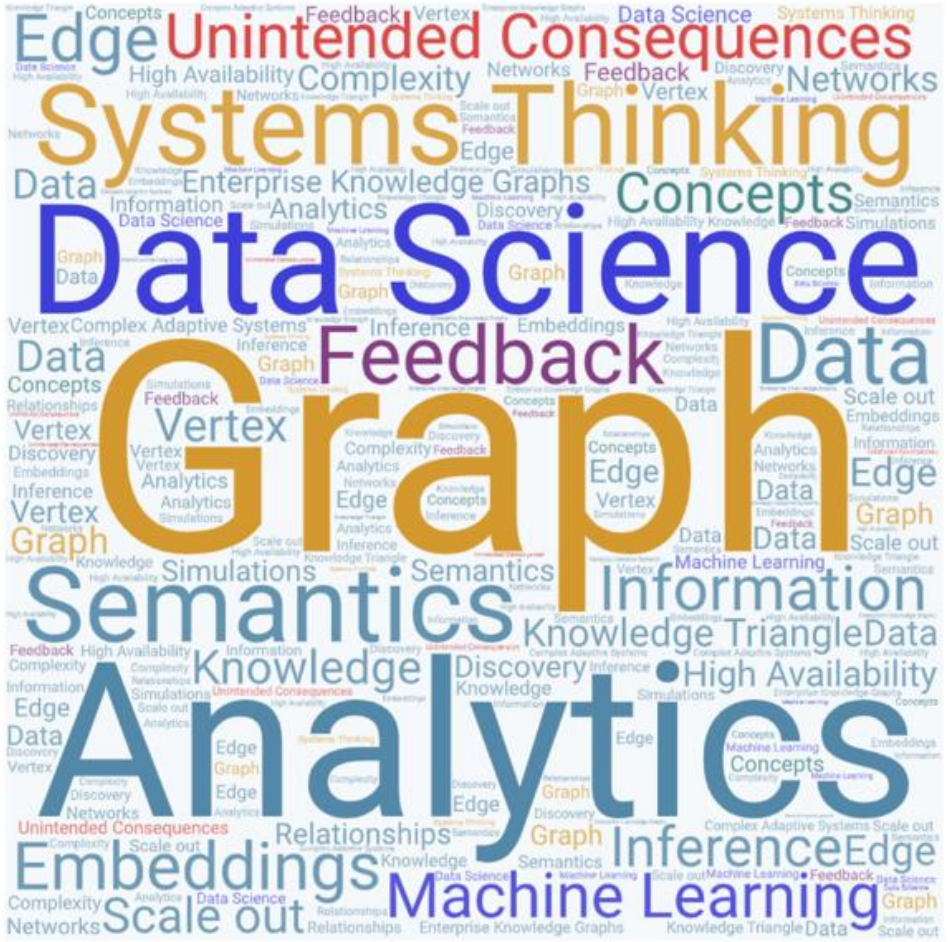
More data creates more precise machine learning models.

Systems Thinking Definition

Systems thinking is a holistic approach to analysis that focuses on the way that a **system's** constituent parts interrelate and how **systems** work **overtime** and within the context of larger **systems**.

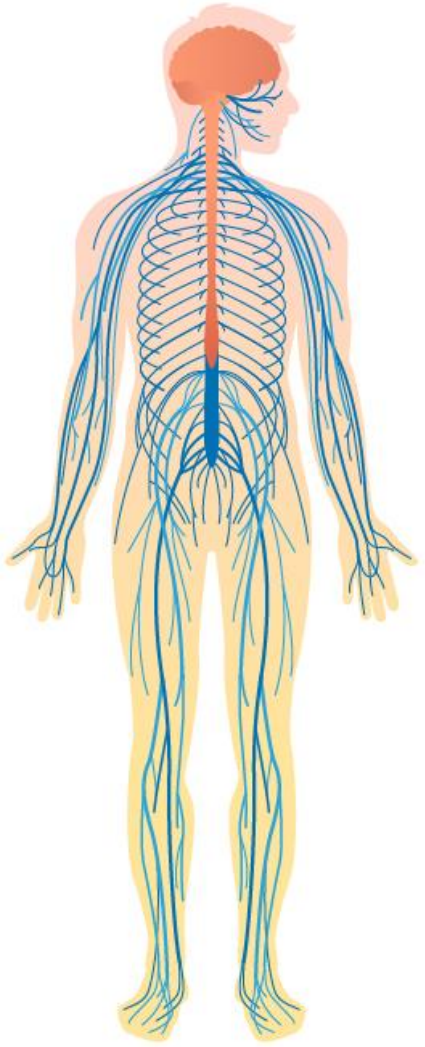
A Definition of Systems Thinking: A Systems Approach
Ross D. Arnold and Jon P. Wade
https://www.sciencedirect.com/science/article/pii/S1877050915002860

A New Discipline: Graph Systems Thinking (GST)



https://dmccreary.github.io/graph-systems-thinking/glossary/

Assumptions

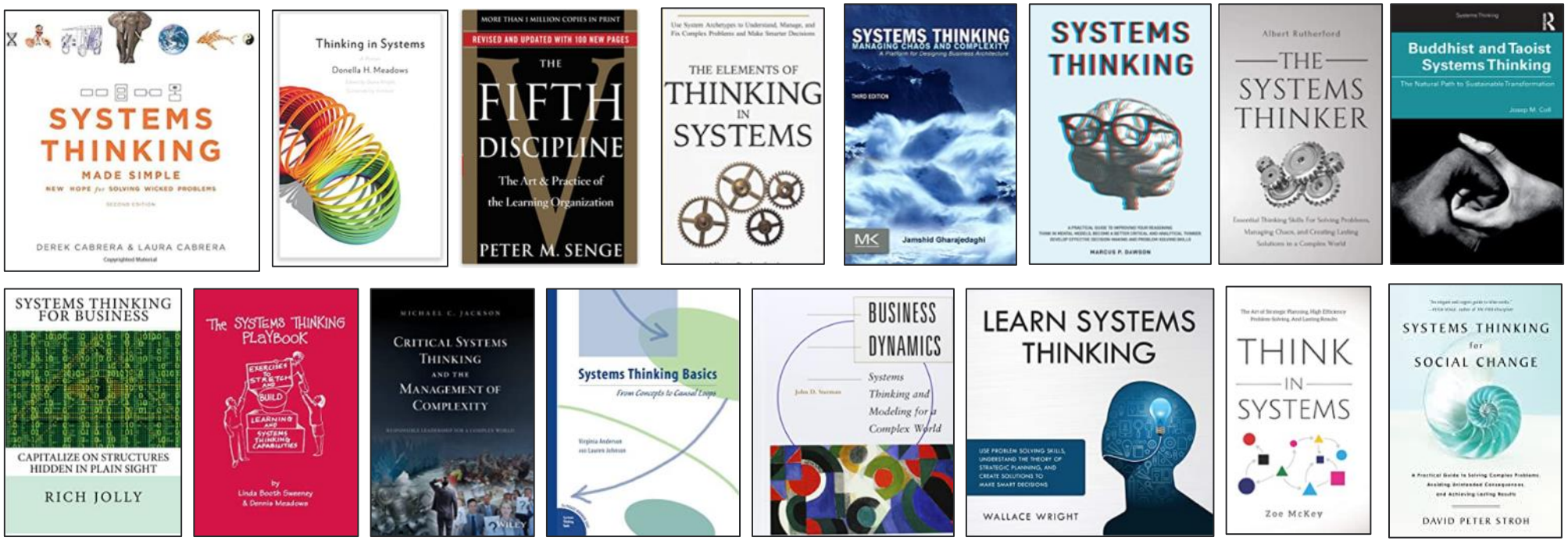


- Enterprise Knowledge Graphs (EKGs) are going to become a central force in organizational dynamics
- They are becoming the **Central Nervous System** (CNS) of organizations
- We need tools to manage the adoption and growth of EKGs
- *Systems Thinking* is an appropriate tool to help us guide EKG Growth

Systems Thinking Terminology

- Archetypes
- Balancing Loop
- Digital Twin
- Dynamics
- Feedback
- Influence Diagram
- Unintended consequences
- Local optimization
- Emergence
- Flow
- Leverage Points
- Limiting Factor
- Nonlinear Relationships
- Resilience
- Reinforcing Loop
- Self Organization
- Sustainability
- Unintended consequences

Many Resources for Learning Systems Thinking



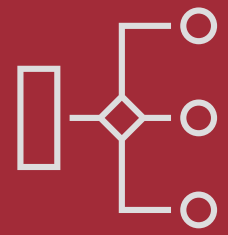
Don’t be intimidated!

Systems Thinking really is about a dozen core concepts

Most people can learn the basic principals of Systems Thinking in a few days

https://dmccreary.github.io/graph-systems-thinking/references/

What is Graph Systems Thinking (GST)?



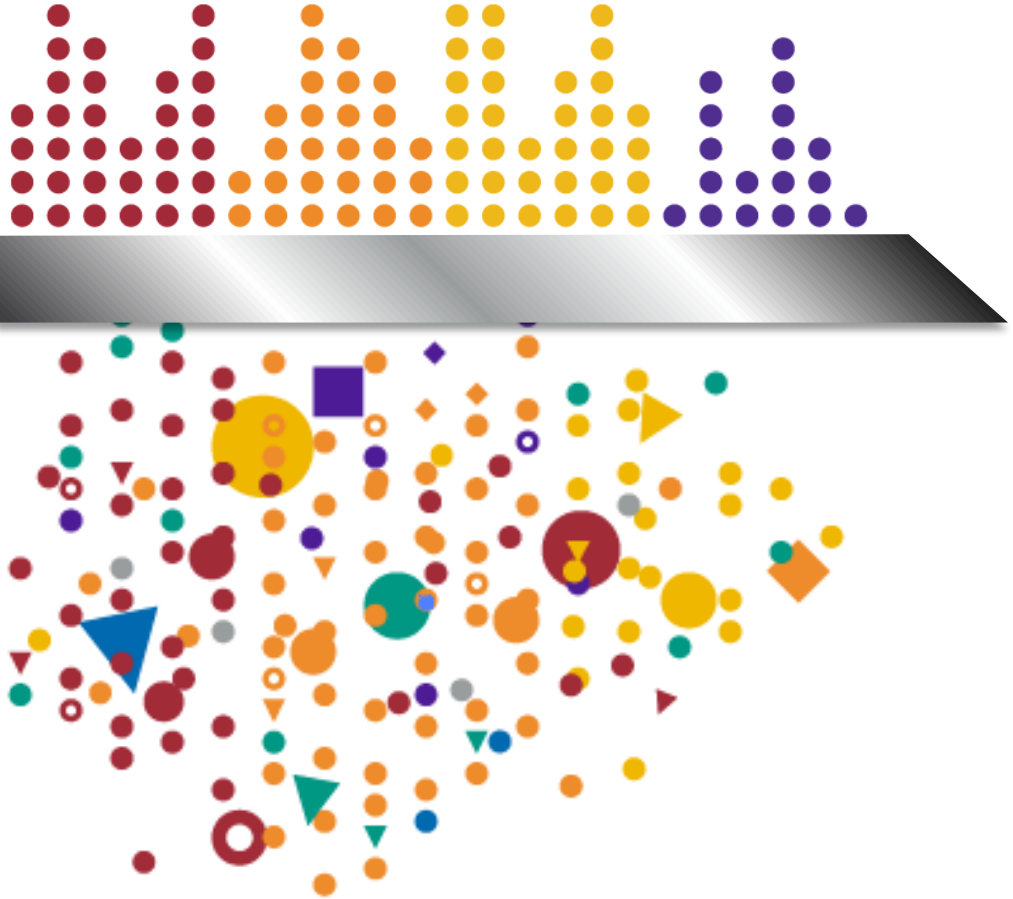
Key Question

How do we choose **what** entities in our organization should be in our EKG?

Answer: Use Systems Thinking!

Edge of Chaos

Consider two regions of your data model:



The EKG Region

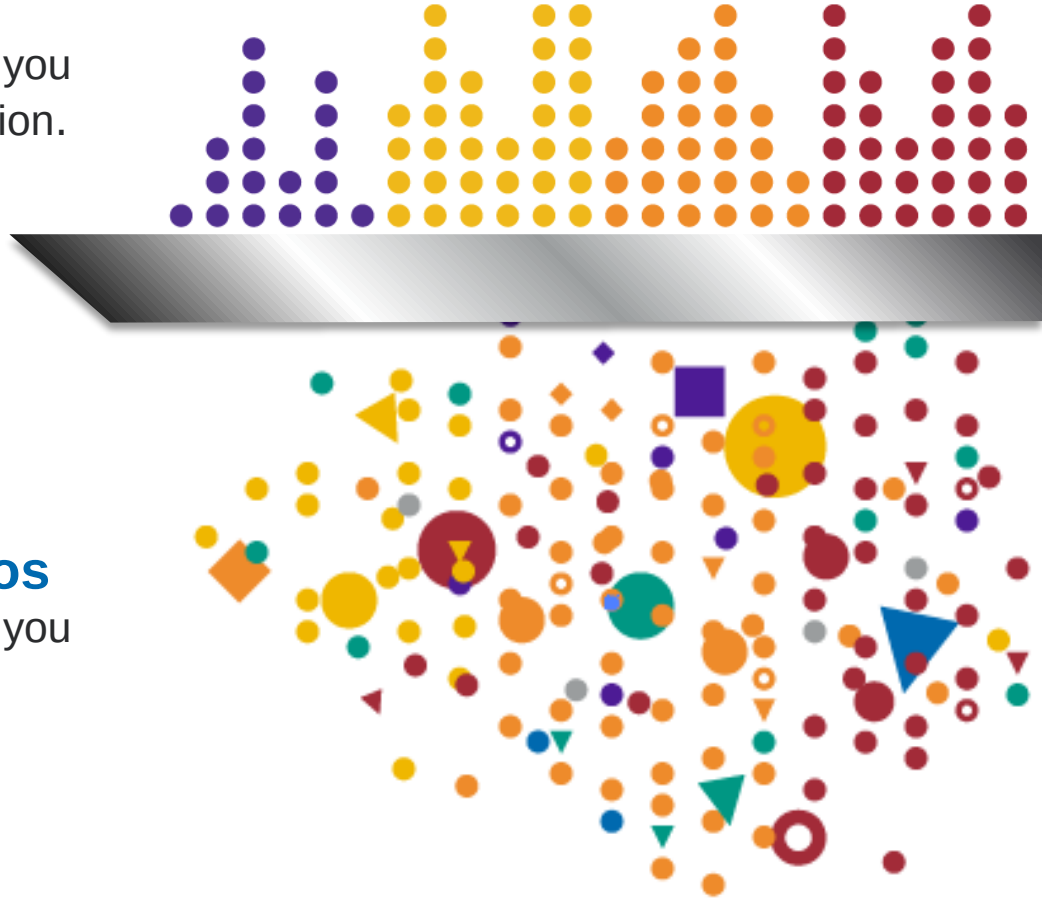
The part of the world that you have modeled with precision.

The Edge of Chaos

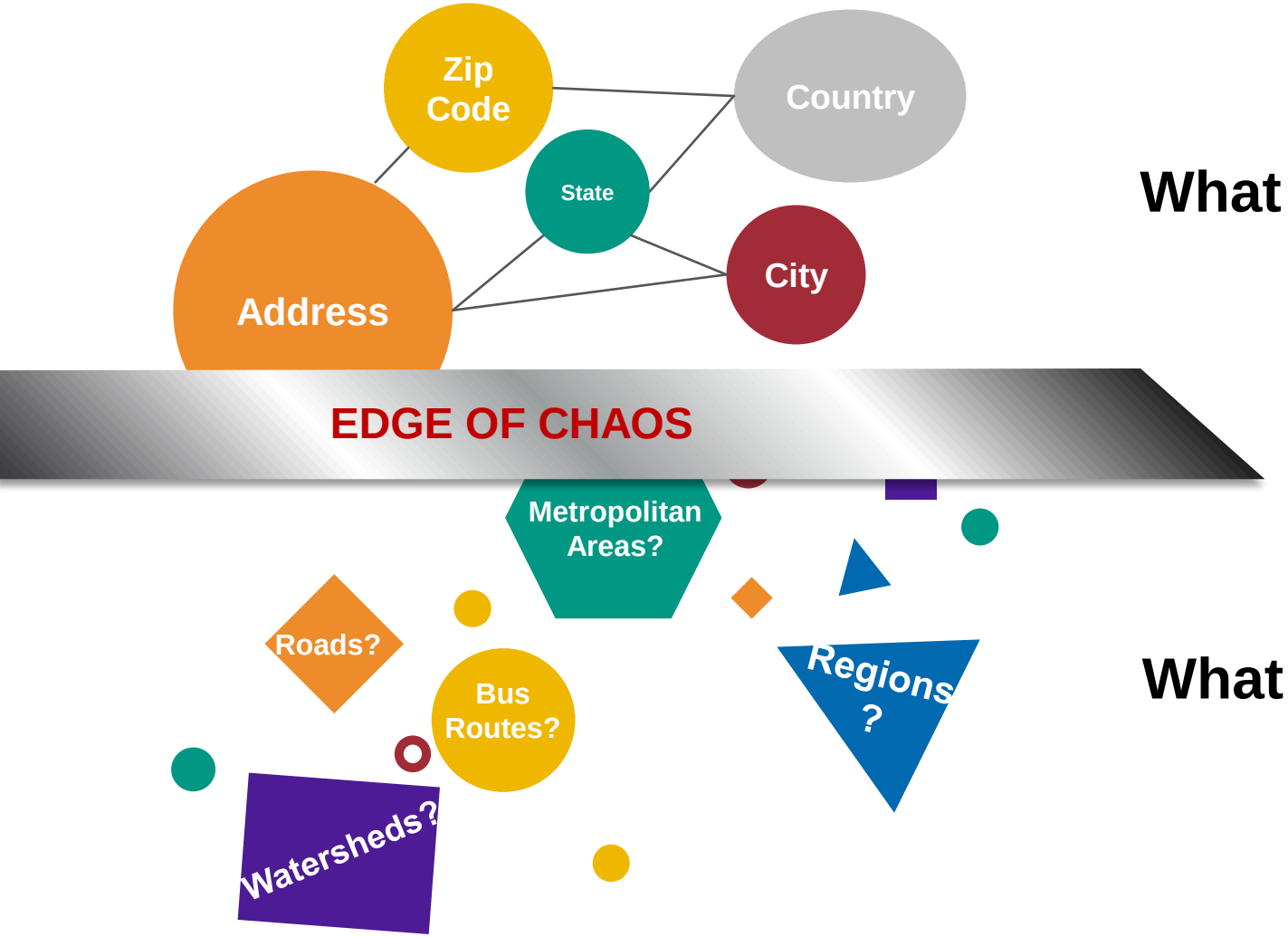
The border between the EKG and the region of chaos.

The Region of Chaos

The part of the world that you have not modeled yet.



Example: Geospatial Models



What you have modeled.

What should you add?

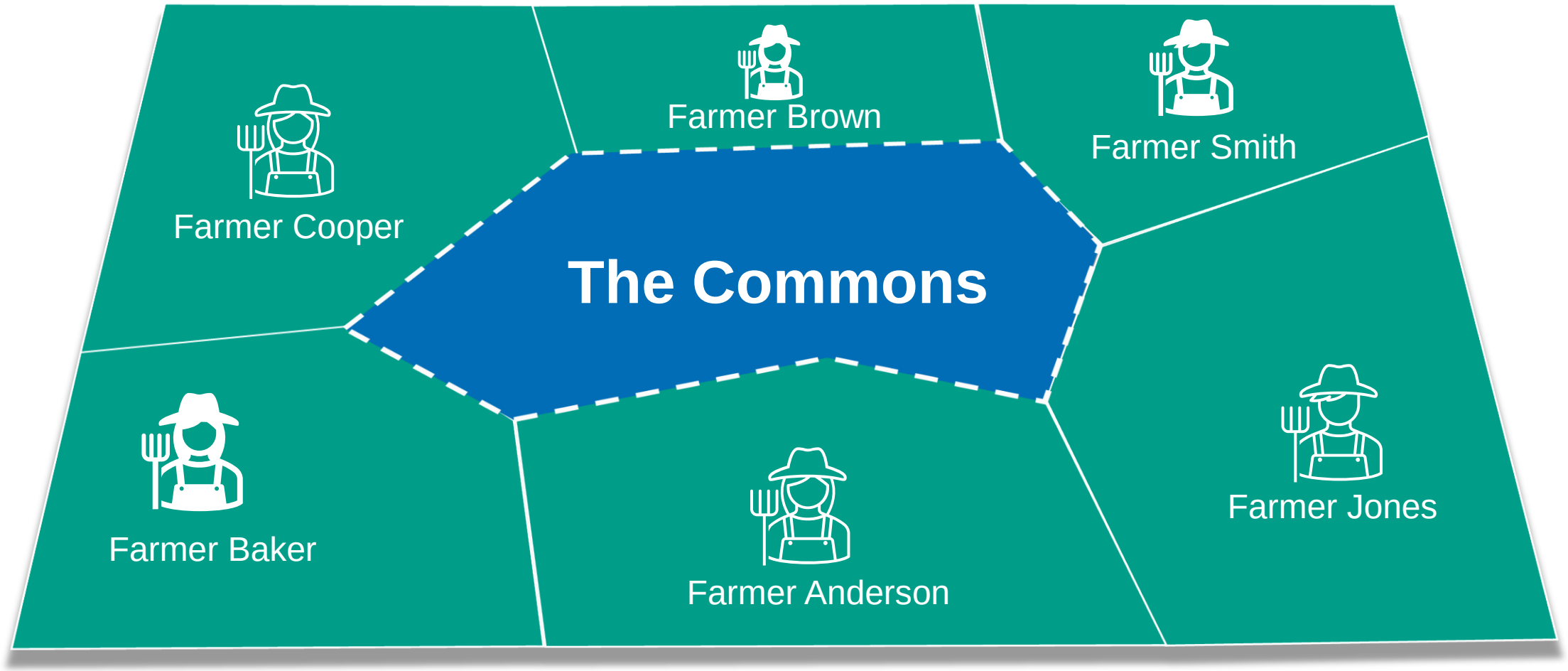
GST Use Cases and Storytelling

OPTUM® | Technology

Example: Provider Recommendation

When our senior members call into our call centers to find a healthcare provider in their area, many of them **don’t drive**. They only want providers that are on **bus routes**. However, our current system does not store this information.

Tragedy of The Commons



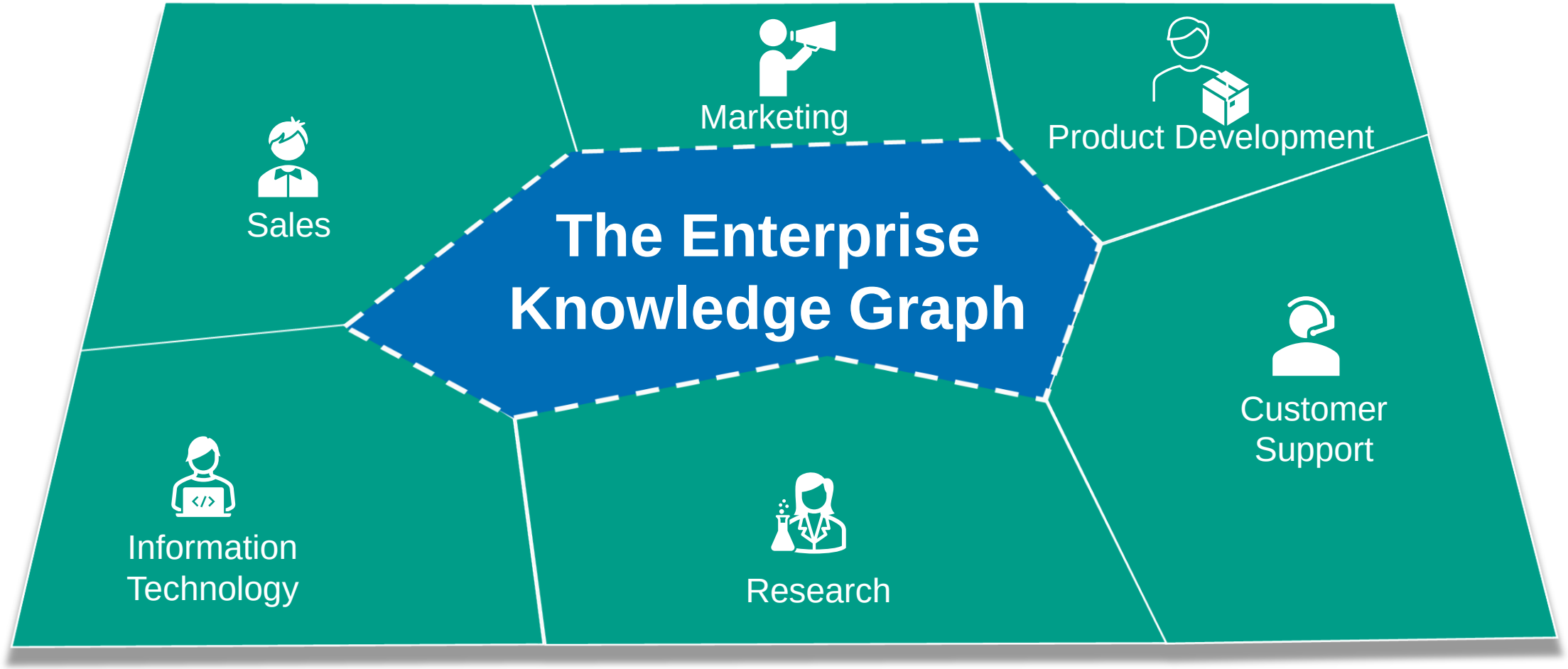
Tragedy of The Commons

A situation in **economic science** when individual users, who have **open access to a resource** unhampered by shared social structures or **formal rules** that govern access and use, act independently according to their own **self-interest** and, contrary to the common good of all users, cause depletion of the resource through their **uncoordinated** action.

https://en.wikipedia.org/wiki/Tragedy_of_the_commons

The more everyone abuses a shared resource, the less valuable it becomes for everyone.

EKGs are Also Shared Resources

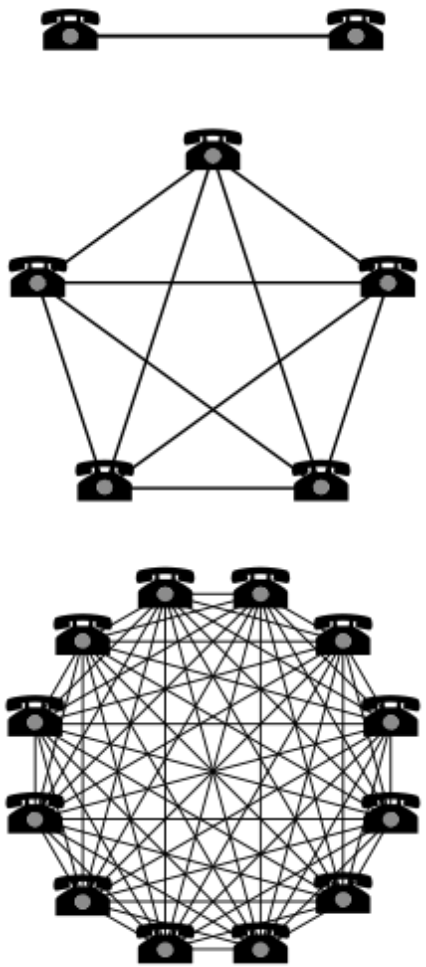


Query Response vs. Time of Day

Average Query Response Time in Seconds (M-F)

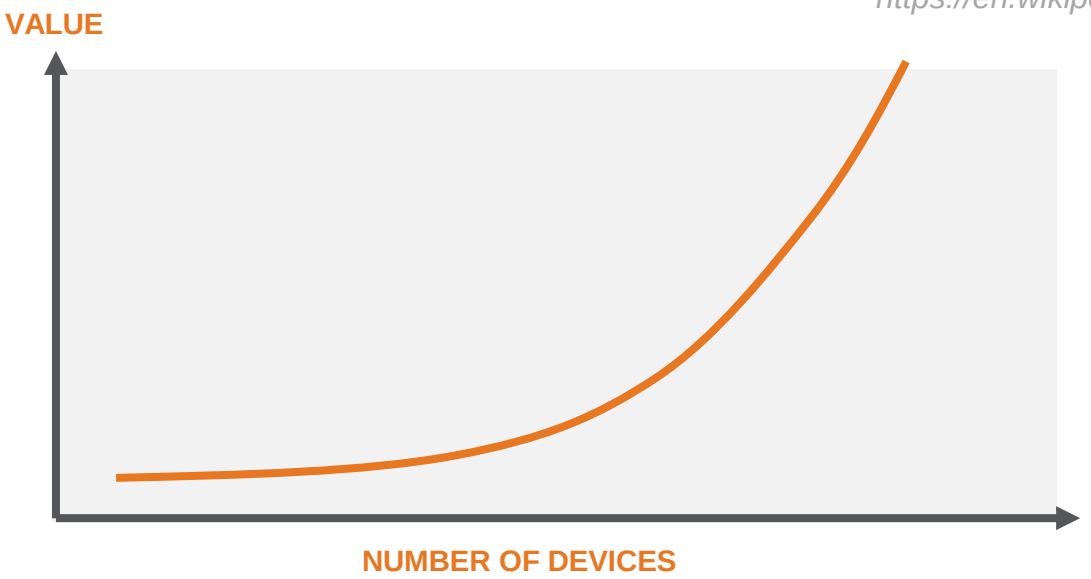
- Many users fighting for a shared resource with limited capability
- The more you use it the lower the value to others

Metcalfe's Law (Reverse Commons Law)



The value of a telecommunications network is proportional to the square of the number of connected users of the system (n^2).

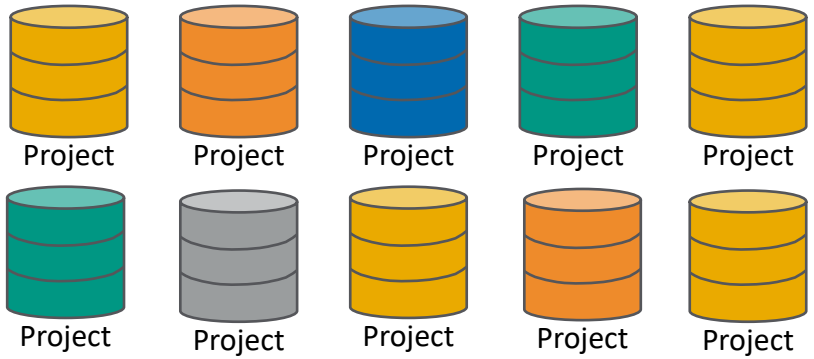
https://en.wikipedia.org/wiki/Metcalfe%27s_law



The more everyone uses a standard, the more valuable it becomes.

OPTUM® | Technology

Project Silos vs Systems Thinking



Isolated Silo View

Each project is an **independent** silo of effort

The success of any project will not impact the success of **other** projects

Project **order** is not relevant and project value is static in time

Project costs and benefits are easy to represent in a simple spreadsheet

The spreadsheet may not reflect the complexities of the real world

Systems Thinking View

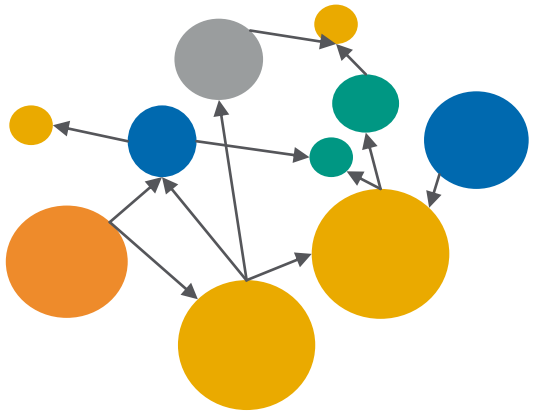
Projects are **dependent** on another project success

The success of **foundational** projects may have a dramatic impact on other projects (x10 faster)

Project **order is relevant** and deferring customer benefit is needed until foundational projects are complete

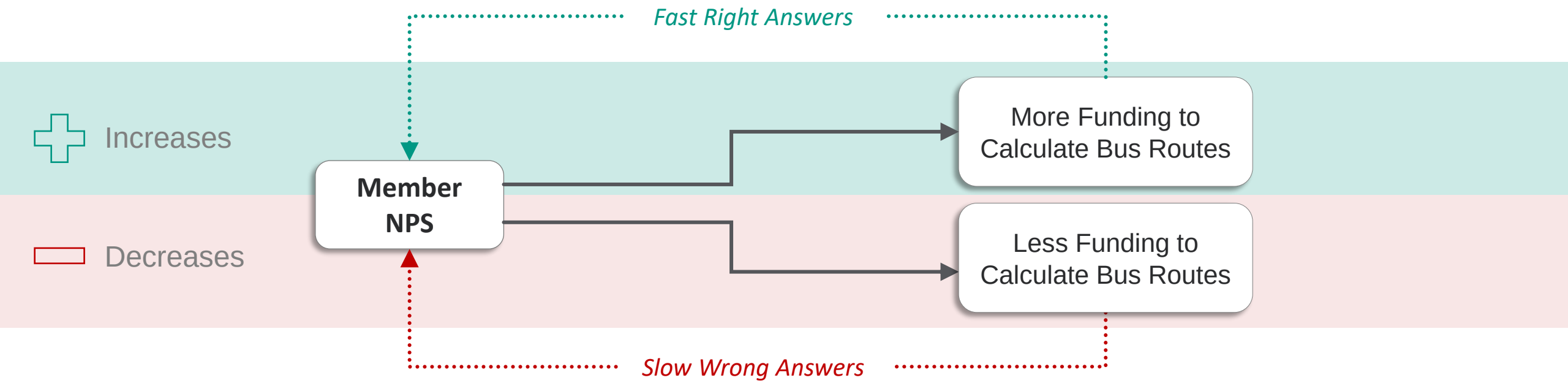
Requires a deep understanding of how resources created by one project can be leveraged by other projects

Reflects the tacit knowledge gained over years of working in research projects and observing different teams' ability to build reusable artifacts



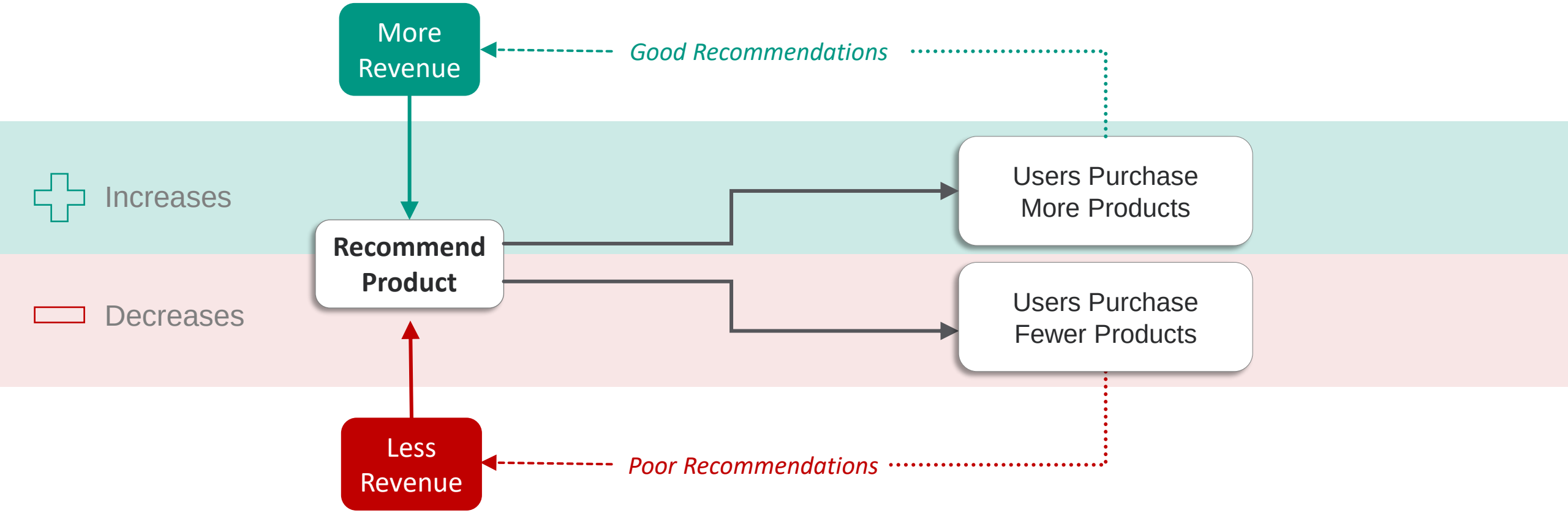
Causal Loop Diagram

Assumption: Higher NPS promotes higher market share

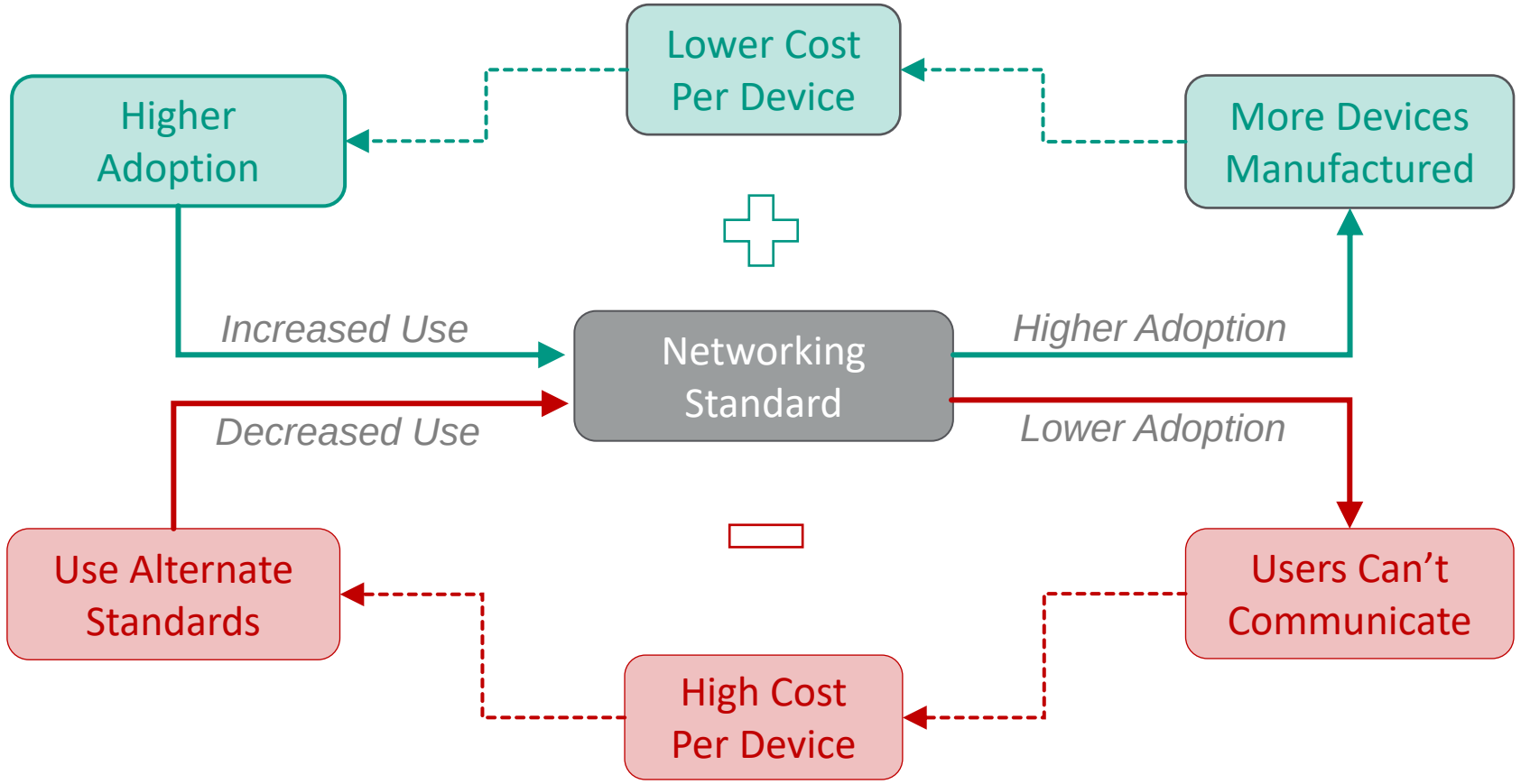


NPS = Net Promoter Score
How happy are our customers?

Predictive Feedback Cycle

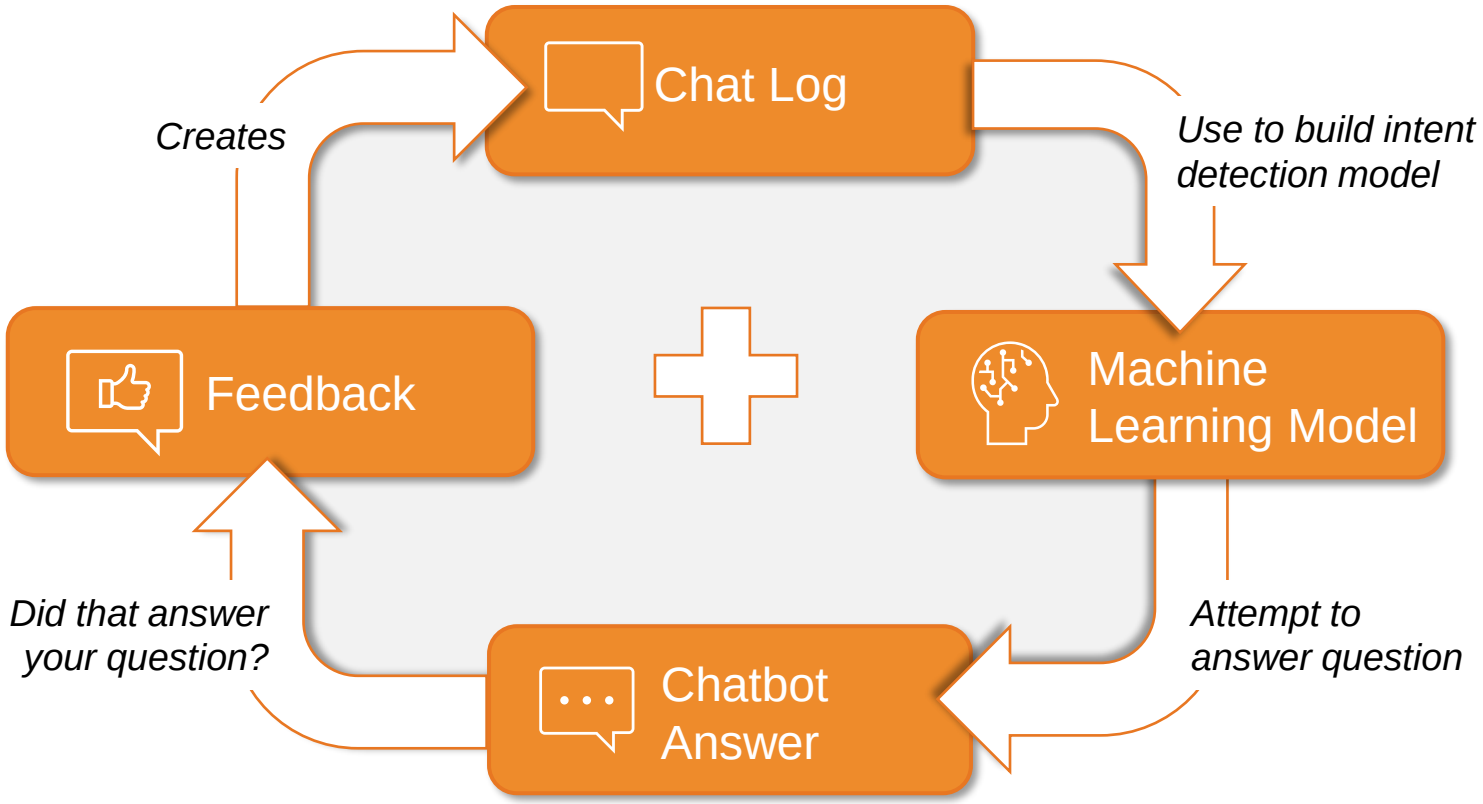


Network Effects (Metcalfe's Law)



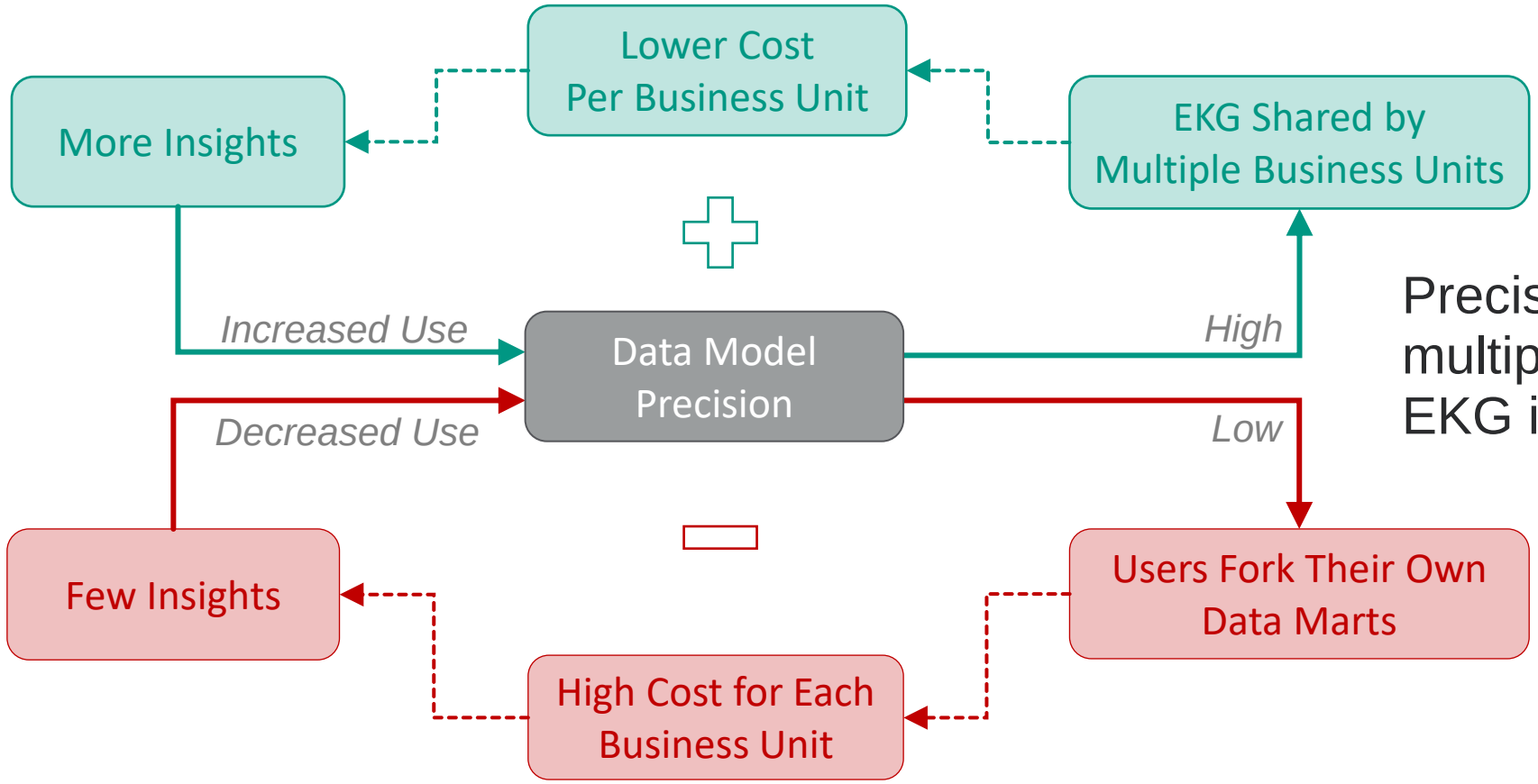
The value of any device on a network standard grows exponentially as the number of connections increase.

Customer Support Chatbot



More feedback is used to build better intent detection models.

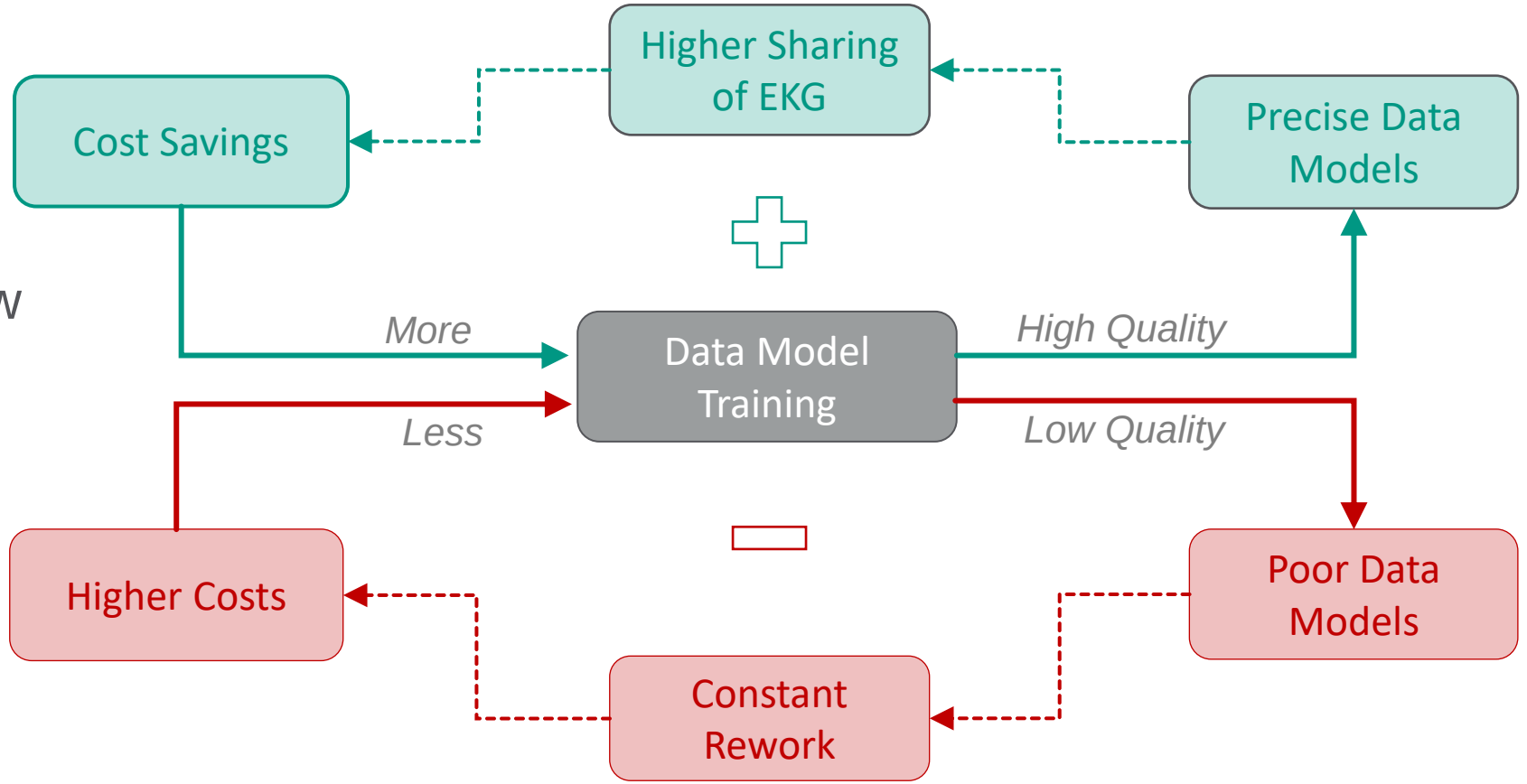
Data Model Precision and Cost Sharing



Precise data models allow multiple business units to share EKG infrastructure costs.

Data Model Precision and Cost Sharing

Precise data models allow multiple business units to share EKG infrastructure costs.

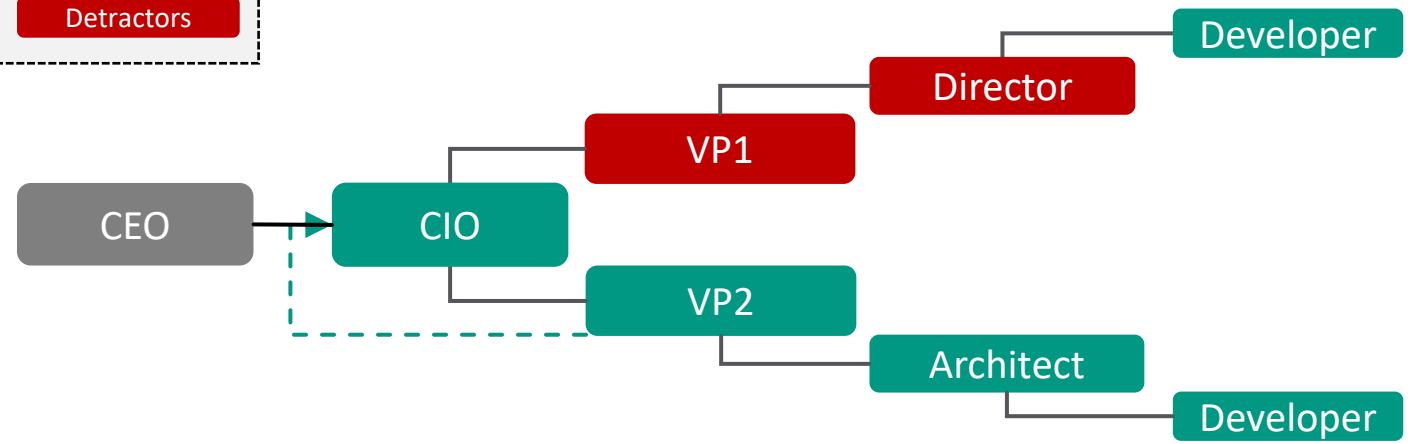


We just modeled this like our RDBMS

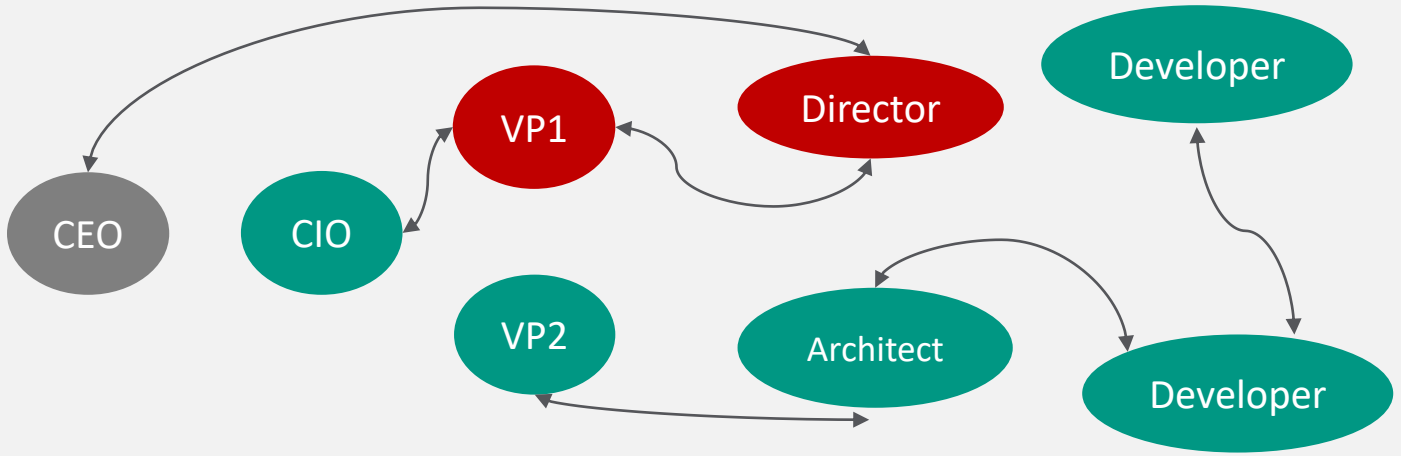
Org Chart vs Influence Diagram

Promoters | Neutral | Detractors

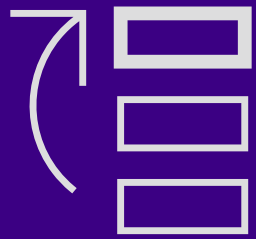
Org Chart Model



Influence Model



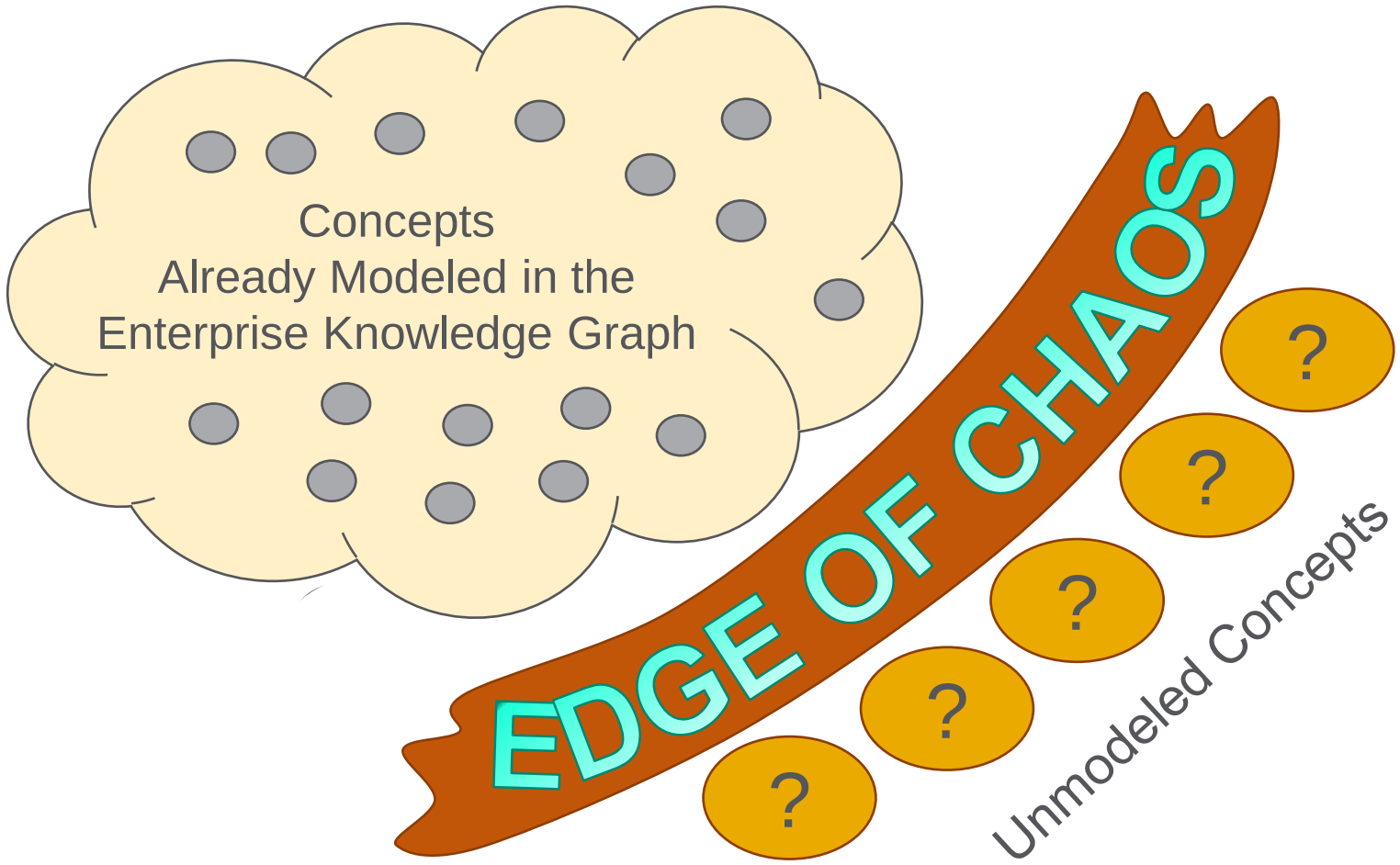
Why is GST Important for Enterprise Architecture?



Connected Data Strategy

Connected Data Strategy: *An **Enterprise Data Strategy** pattern that brings **focus** to the business **value** of connecting disparate silos of data.*

Edge of Chaos



Consider two regions of your data model:

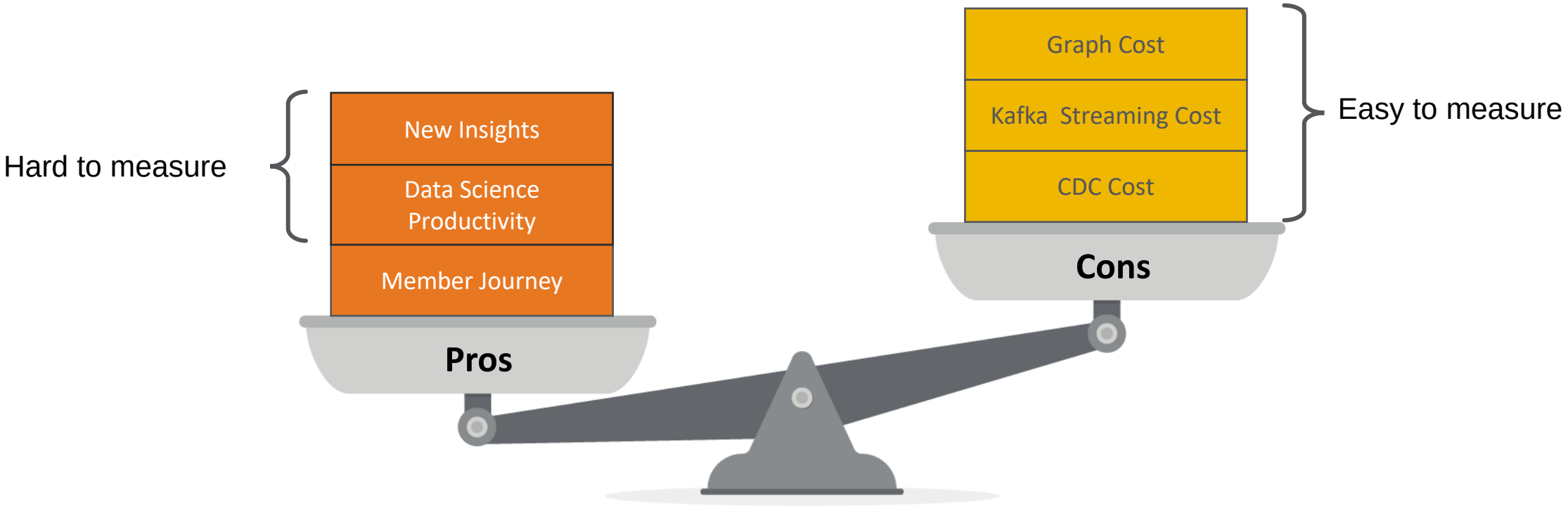
1) The part of the world that you have modeled with precision. We call this the known concepts region.

2) The part of the world that you have not modeled yet. We call this the “region of chaos”.

The border between the EKG and the region of chaos we will call “The Edge of Chaos”

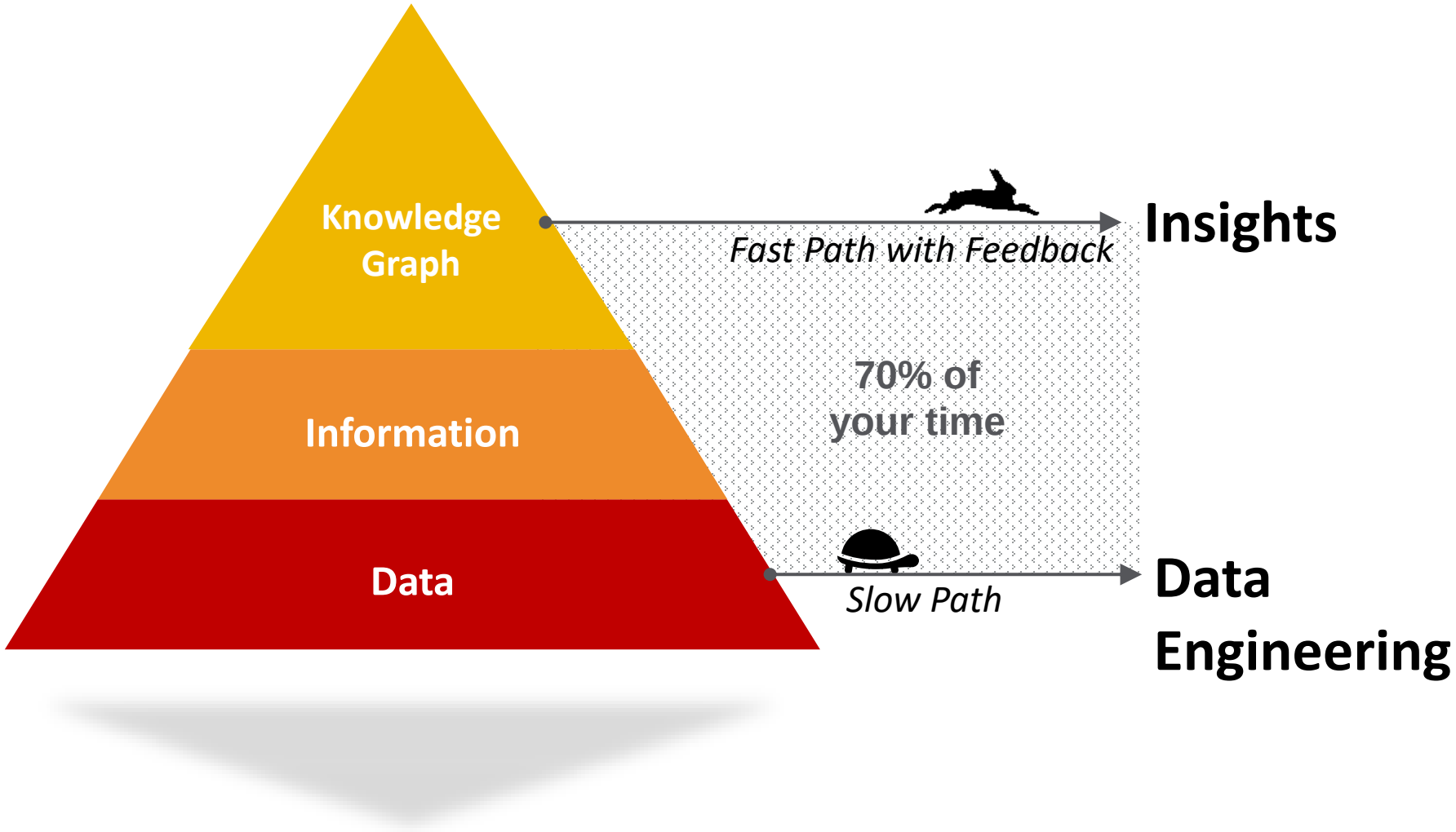
Systems Thinking helps us understand how to grow our EKG

Goal: Objective Weighing of Pros and Cons



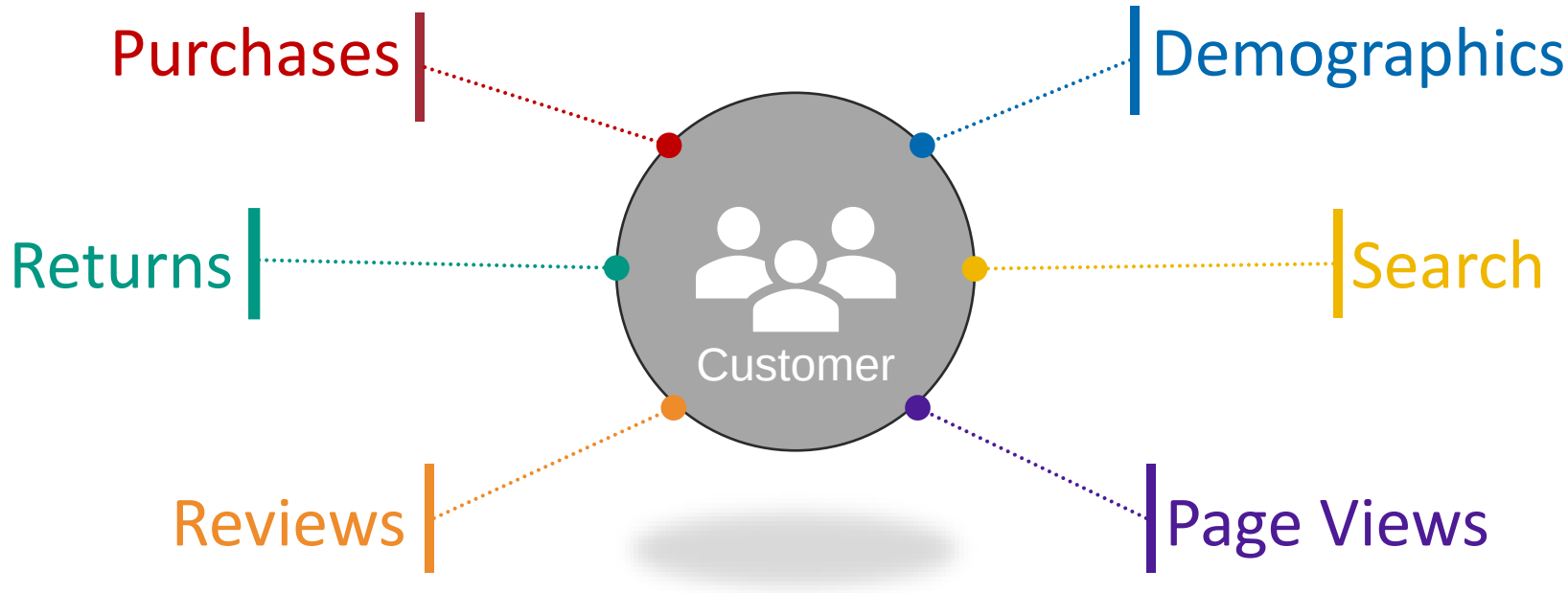
- We are starting to learn the benefits of connected data: Customer 360
- It takes time and effort to create high-quality connected data
- Can we objectively measure each of the pros and cons?
- **What is the value of "new insights?"**
- How is our decision making driven by what is easy to measure?

From Data Scientist to Knowledge Scientist



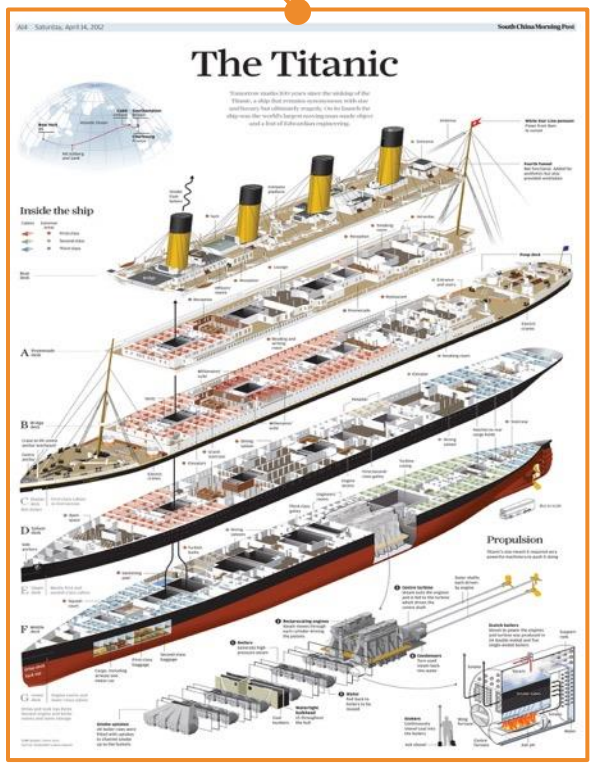
Customer-Centric EKG Strategy

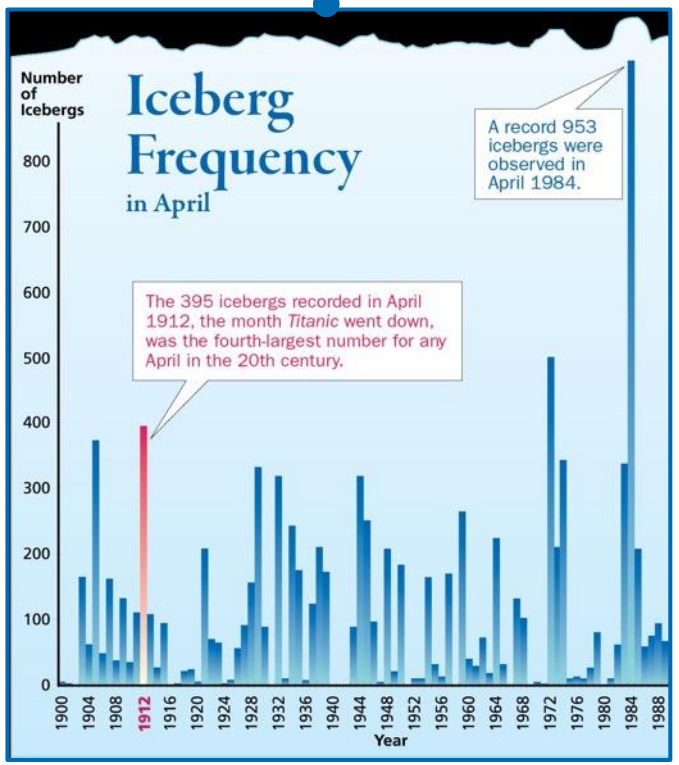
Put your customer at the center of your EKG



Modeling Precision

"All models are wrong. Some models are useful." - George Box





Highly normalized graph data models are better at modeling the real world than rigid denormalized tabular data

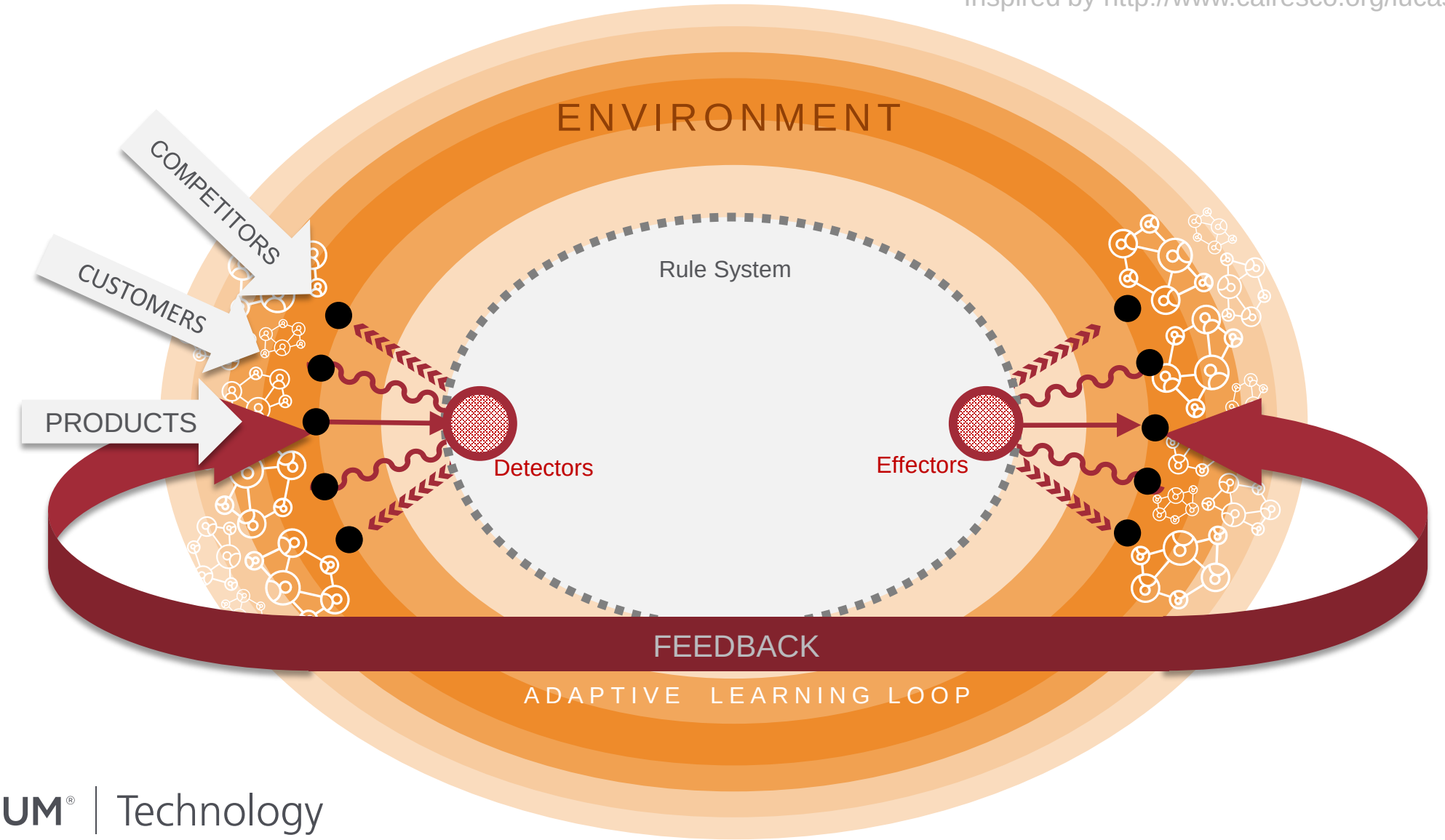
Models that are correct are easily shared across business units

The more business units that share a model, the less they duplicate data and the lower the cost of storing and analyzing the data

Conclusion, Request for Help, and Recommended Next Steps

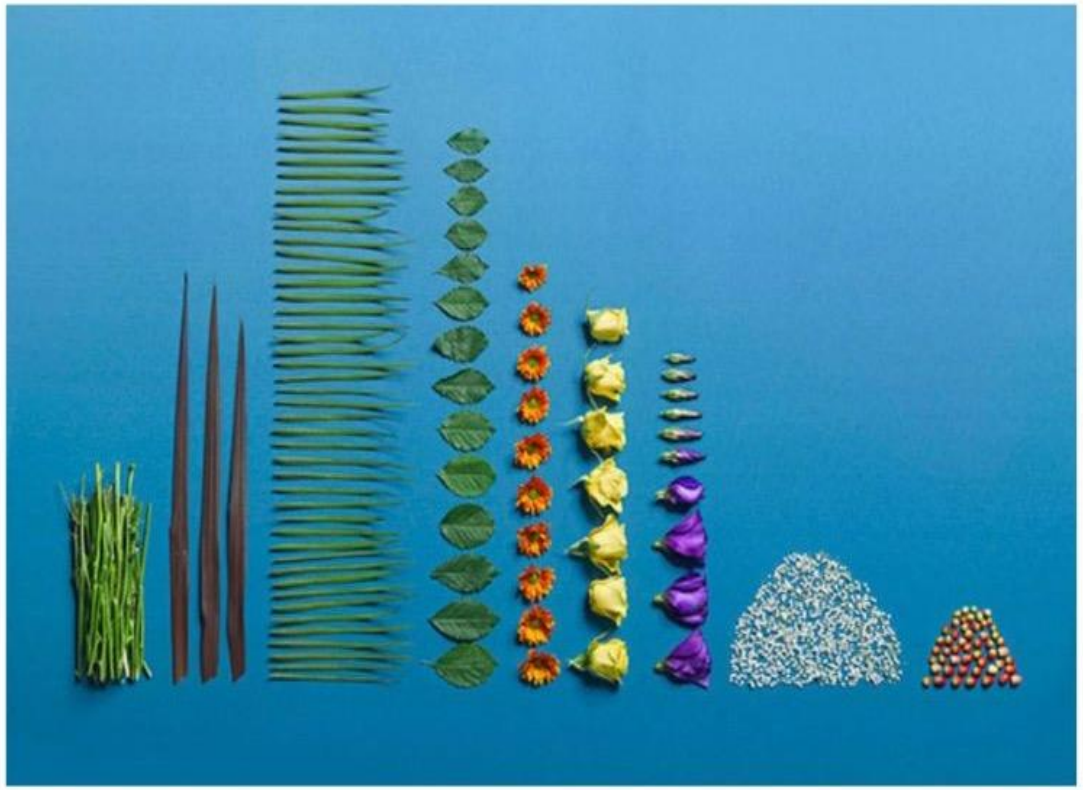
Complex Adaptive System As Rules

Inspired by http://www.calresco.org/lucas/cas.htm



Graphs

Relational Databases



The Art of Clean Up: Life Made Neat and Tidy by Ursus Wehrli

Graph Databases

Request for Help

Please share your architypes!

Architypes are the “Patten Library” of GST (think Design Patterns)

Open-Source Repository of GST Architypes: https://github.com/dmccreary/graph-systems-thinking

Please help other Enterprise Architects reuse these stories

Recommendations

1. Study Systems Thinking
   Many online resources
   A great tool for enterprise architects
2. Promote Systems Thinking in your organization
   Invite other great systems thinkers to speak
3. Apply Systems Thinking as a tool to create more **connected** data
4. Stop designing to minimize JOINs. Design to maximize graph model sharing.
5. Use Systems Thinking to **justify** and **grow** your EKGs!

Thank You! My Three “Systems Thinking Gurus”

Arun Batchu
Gartner
Systems Thinking Guru

Nikhil Deshpande
Intel PIUMA
Hardware Graph Guru

Kumar Deepak
Xilinx
FPGA Guru

Thank you!

Dan.McCreary@optum.com

https://www.linkedin.com/in/danmccreary

https://dmccreary.medium.com/

@dmccreary

OPTUM