



The Intersection of Learning, Knowledge and Language

Graph + AI Conference
October 5th, 2021

Dan McCreary
Distinguished Engineer



OPTUM[®]

Advanced Technology Collaborative

Talk Description

Title: The Intersection of Learning, Knowledge and Language

Description: In the past ten years, we have seen a revolution in machine learning. It has enabled deep insights in finding patterns in images and language in real-time.

Now we are on the verge of a new revolution, the rise of custom graph hardware and graph machine learning.

This session will show how Graph Systems Thinking (GST) and parallel computation drive innovation in managing our enterprise knowledge graphs. We move from the superficial world of cosine similarity and recommendation systems to machine learning to create graph embeddings. These trends will accelerate real-time insights in highly connected data and give us a clear path to the proper training to keep our organizations globally competitive.

Enterprise Knowledge Graphs (EKG) Are Changing the Way We Think

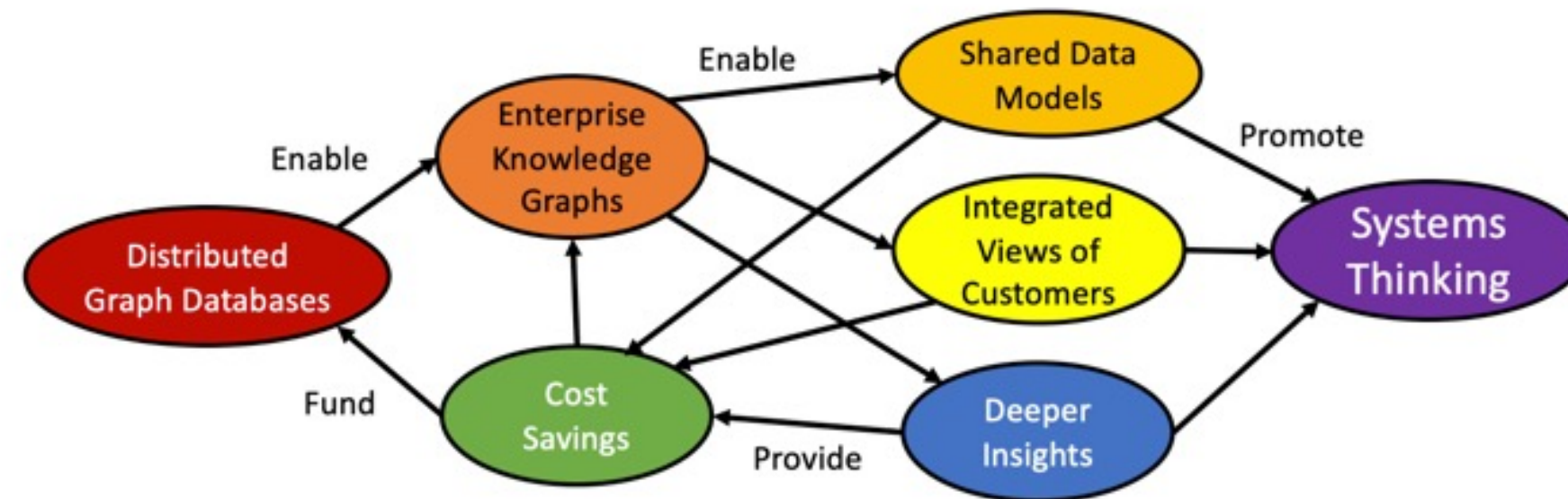
The God Graph



Dan McCreary Nov 1, 2020 · 5 min read



How EKGs and Systems Thinking can make the world better

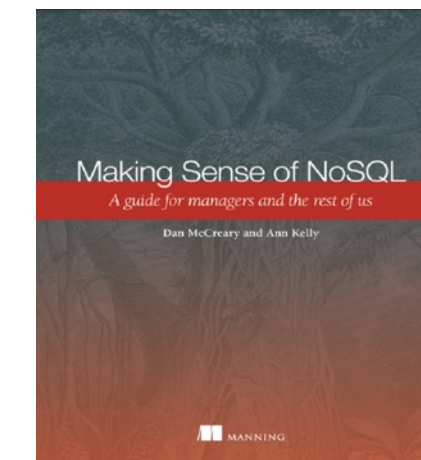


How distributed graph databases are encouraging Systems Thinking, which is making the world a better place.

Blog post on Medium on November 1st, 2020

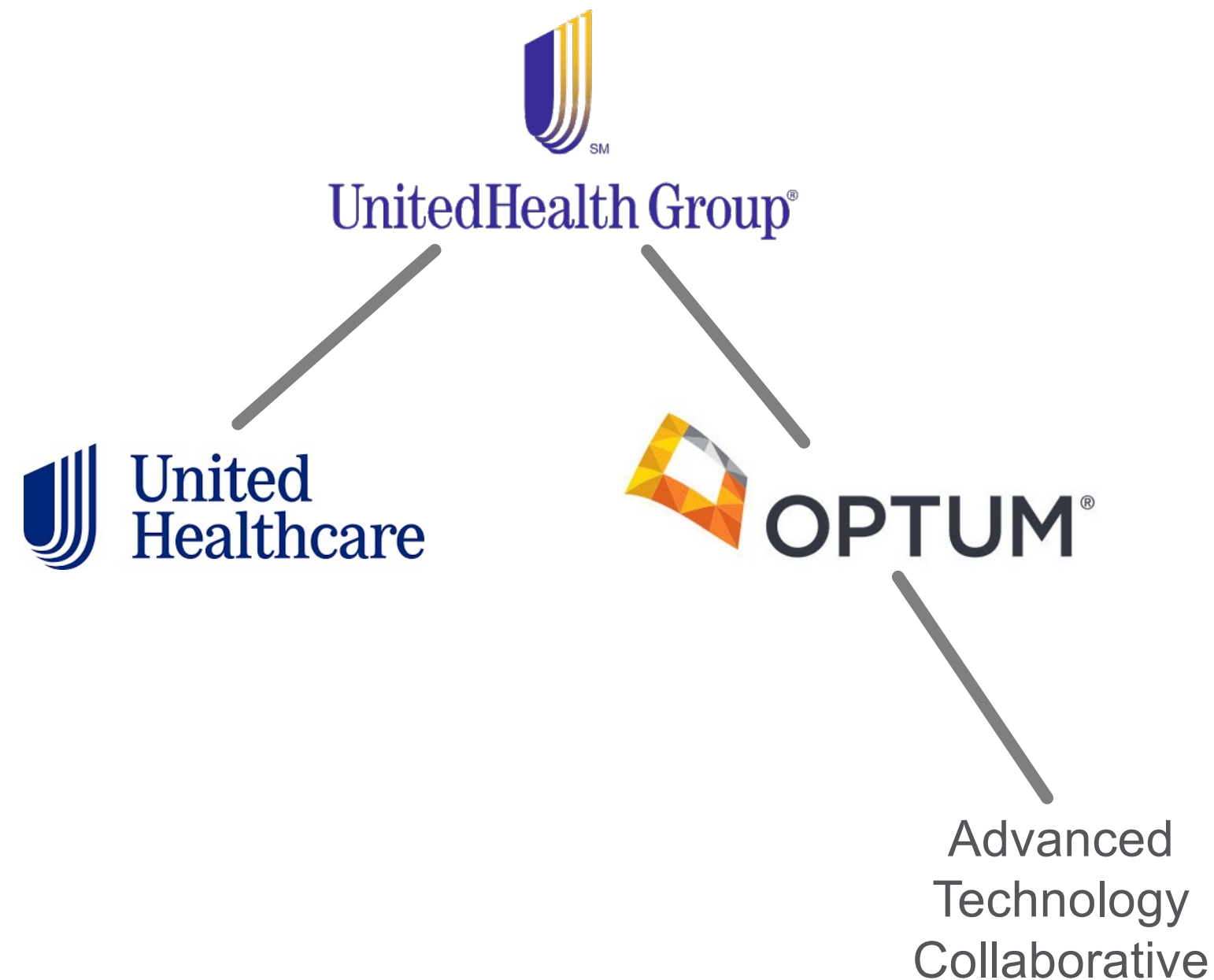
Hello, my name is Dan

- Distinguished Engineer in AI and **Graph Technologies**
- Co-founder of "NoSQL Now!" conference
- Co-author of "*Making Sense of NoSQL*" (w. Ann Kelly) and "*State of Healthcare Technology*"
- 20+ years of working with NoSQL
- Background in solution architecture, knowledge representation for AI, data strategy, NLP, mentoring and curriculum design for technology solutions
- STEM curriculum volunteer



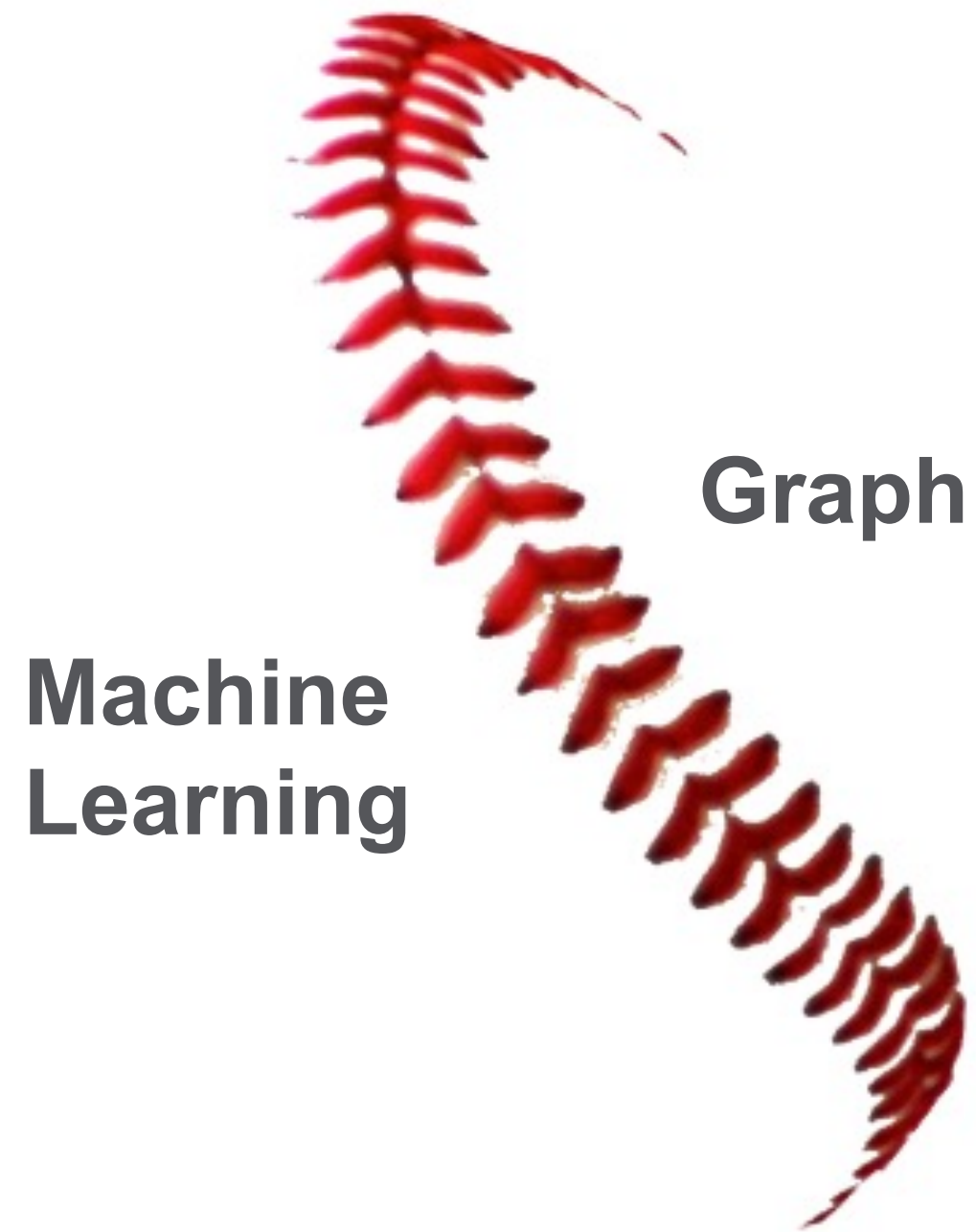
Personal Mission: Help organizations objectively match business problems to emerging technology with **storytelling**

About Optum and the Advanced Technology Collaborative



- UHG is a Fortune 6 firm with over 330,000 employees
- United Healthcare sells healthcare insurance
- **Optum** provides technological, operational and consulting solutions and services to healthcare organizations, pharmaceutical companies as well as the federal and state governments
- Optum has:
 - 35,000 **IT** employees
 - 3,200 data scientists
- The mission of the Optum ATC is to evaluate emerging technologies and apply them to business challenges through pilot projects

Innovation Occurs at the **Seams** of Technology Domains

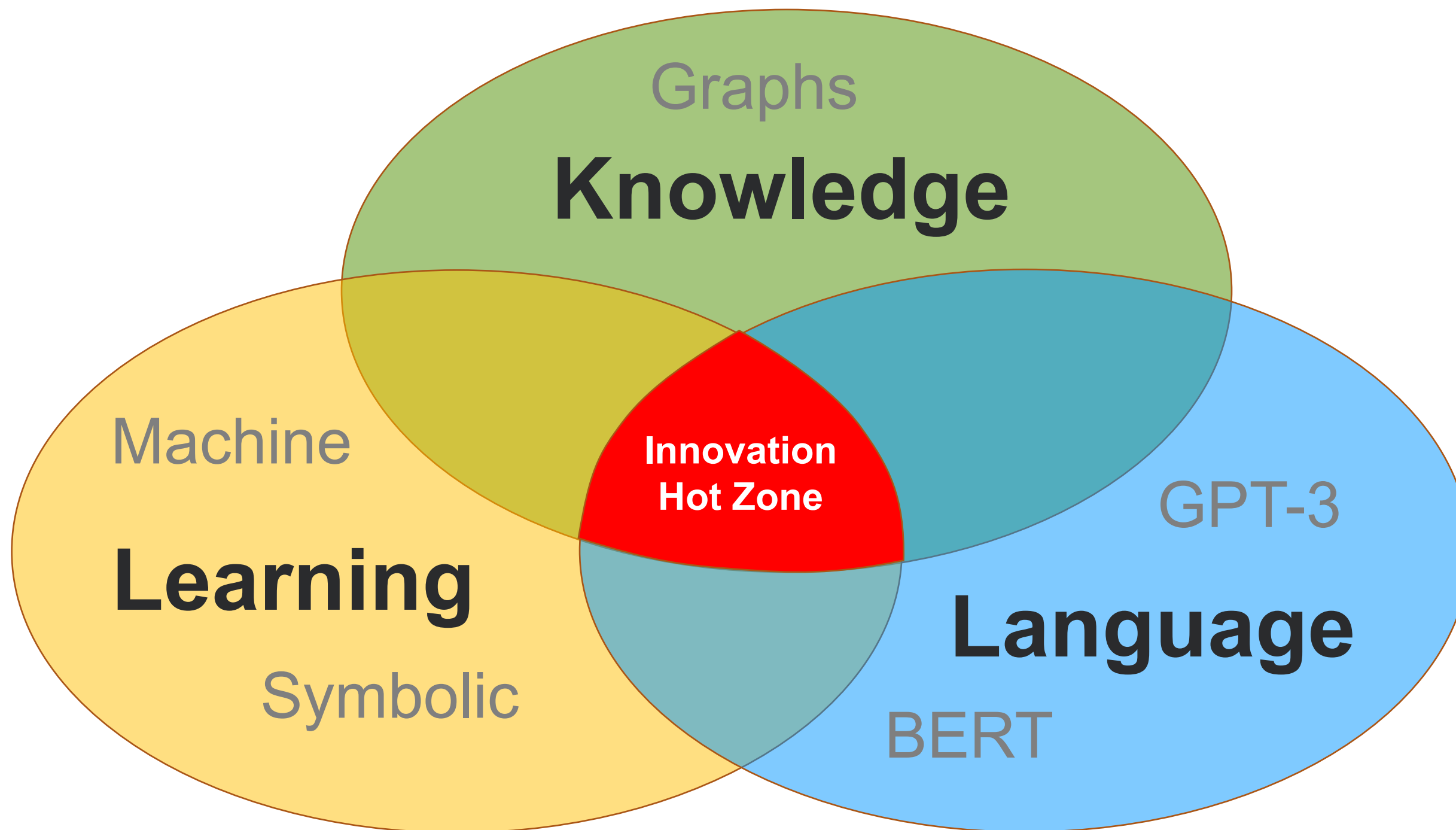


As a technical generalist I have found that tremendous innovation occurs at the **seams** of technology domains.

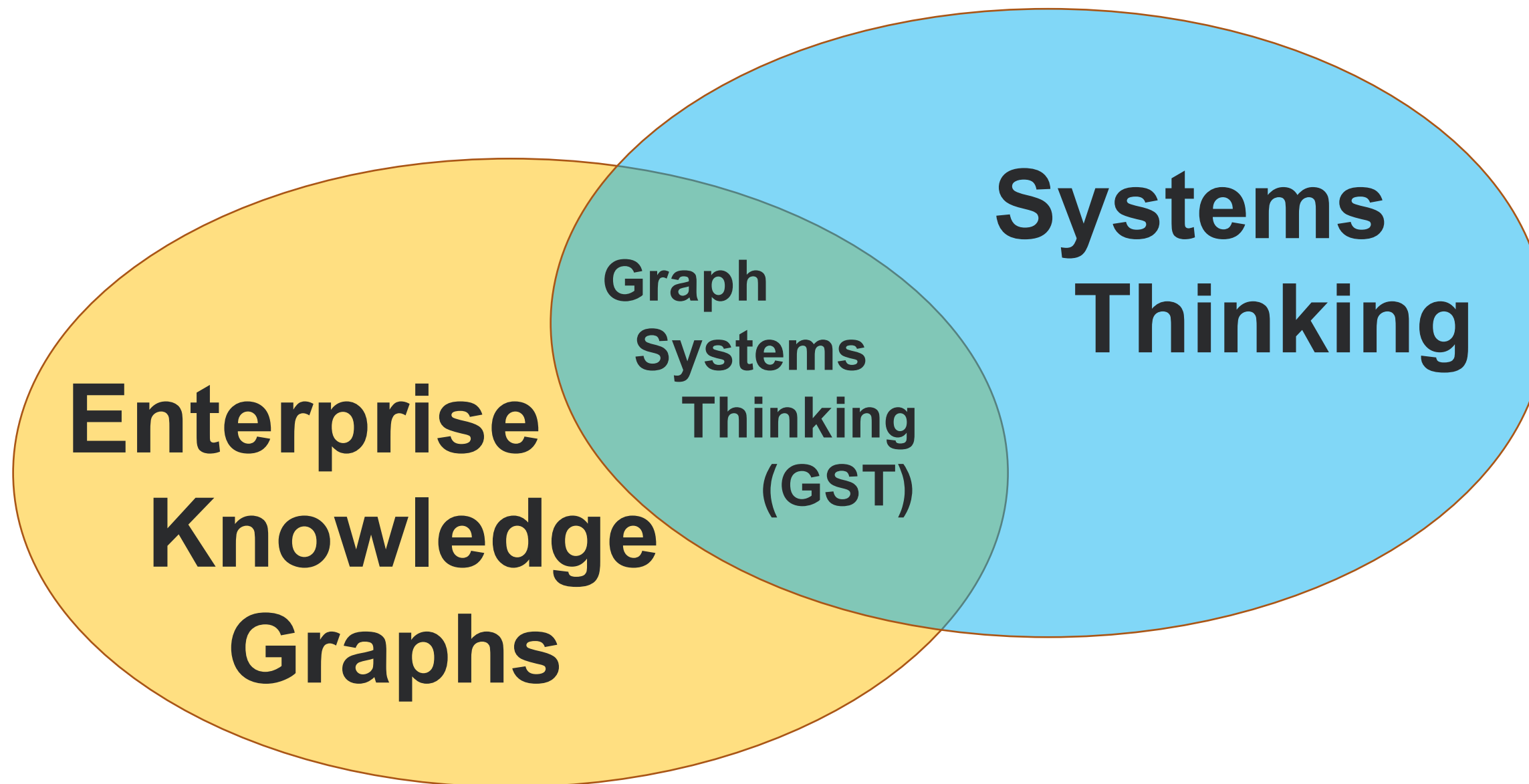
– David Kanter (MLCommons)



The “LKL: Innovation Hot Zone



Enterprise Knowledge Graphs and Systems Thinking

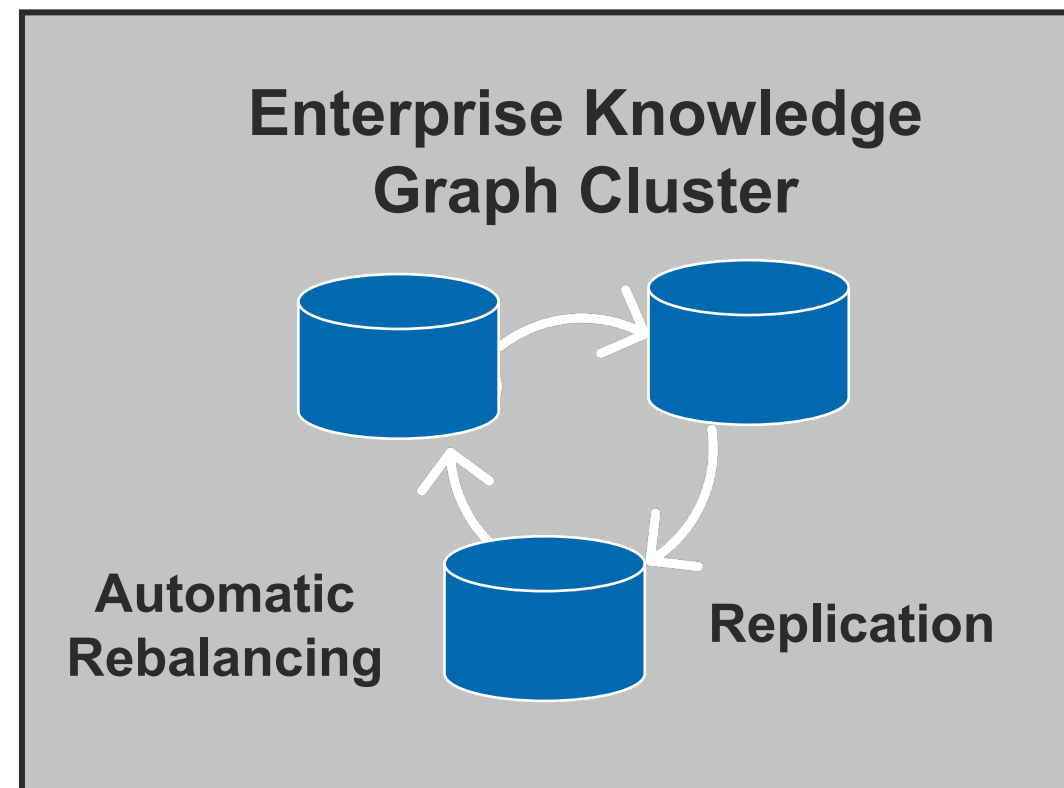


What is an Enterprise Knowledge Graph (EKG)?



Hundreds of
Applications

Secure
Ad-Hoc
Query



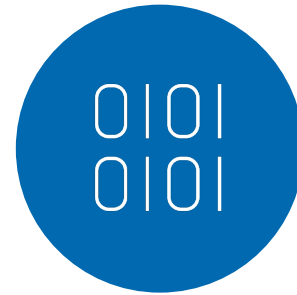
- High Availability (HA): over 99.99% uptime
- Rolling upgrades: never shut down services for upgrades
- Vertex-level Role-based Access Control (RBAC)
- Resource quotas
- Large library of graph algorithms

Definition: 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 Characteristics of an EKG



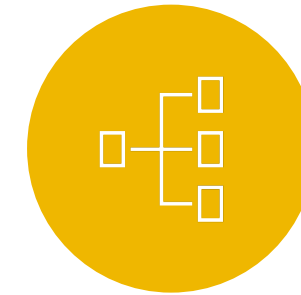
**Scale-out
data size**



**Scale-out
compute**



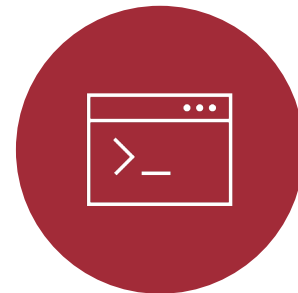
**Scale-out
security**



**Scale-out
manageability**



**Scale-out
data quality**



**Scale-out
algorithms**



**Scale-out
query**

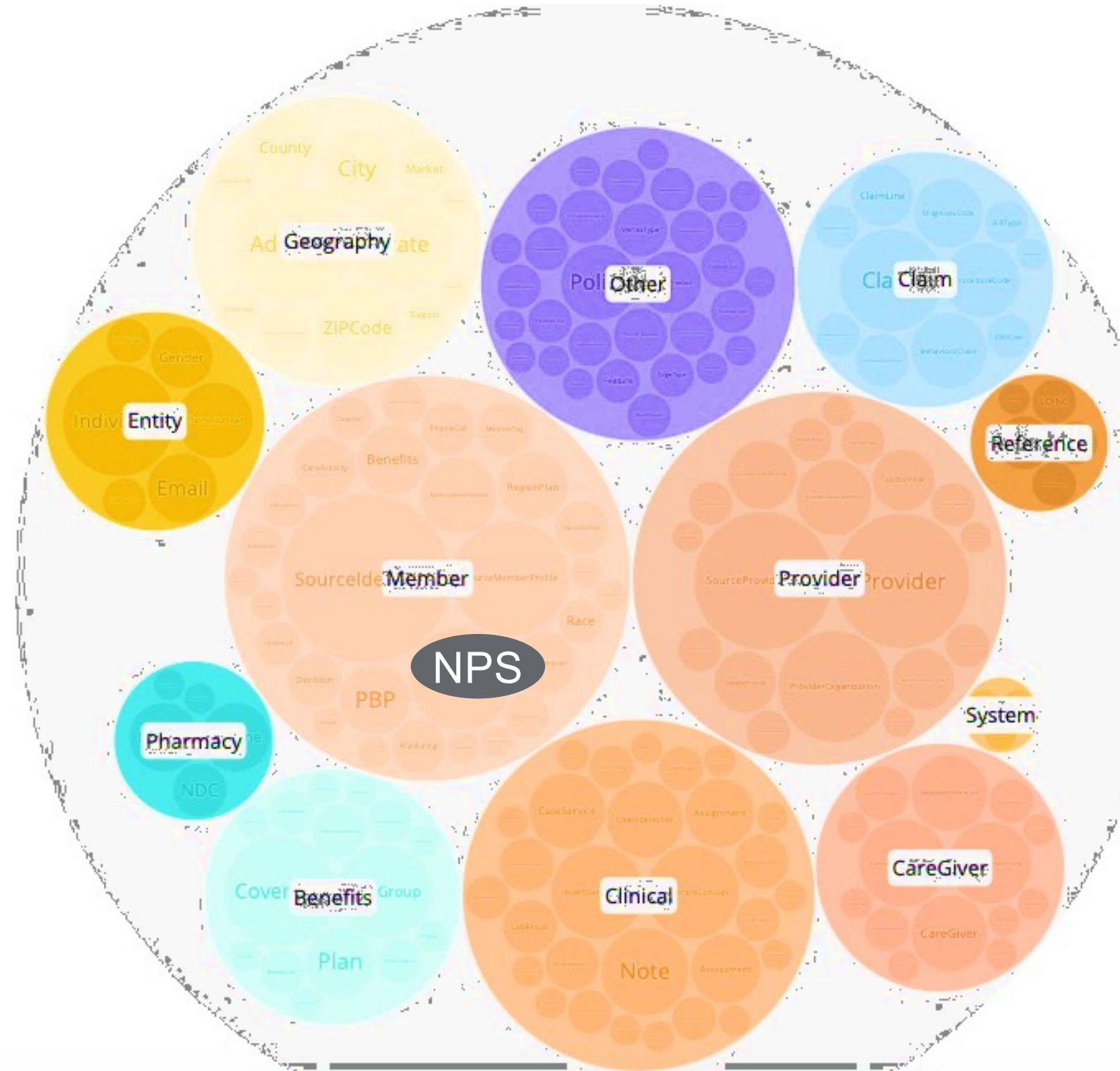
Beware of false prophets!

UHG/Optum Enterprise Healthcare Graph

- Billions of vertices
- 10s of Billions of edges
- Tens of millions of updates per day
- Streaming interfaces: 60 seconds from change in operational source system to updates in the Healthcare Graph
- 25K concurrent users with 100msec response times and 99.99% uptime
- Deep insights in clinical value

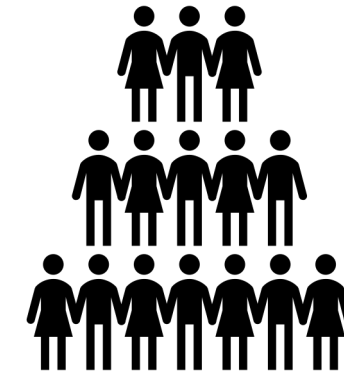
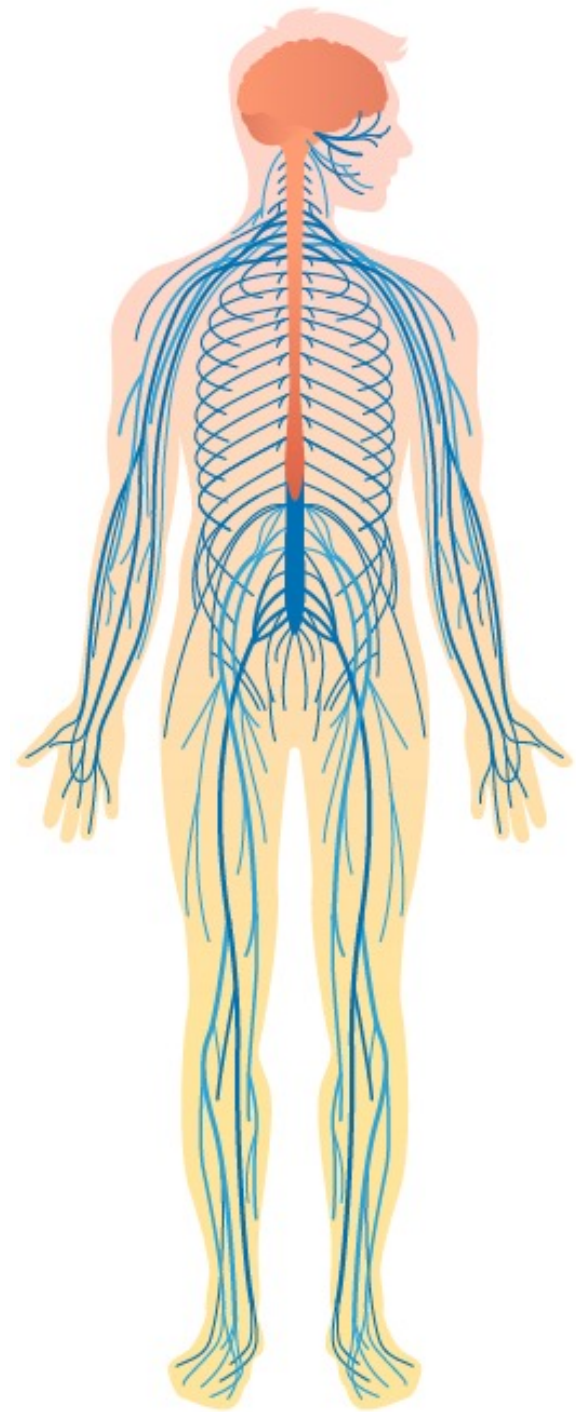


Cross Discipline Queries



- 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 and clinical claims C ?
- Are the results supported by **real-world evidence**?

Goal: Become the Central Nervous System for Healthcare



1. Care coordinators
2. Physicians
3. Providers
4. Nurses
5. Call center agent
6. Senior living assistants

Get the:

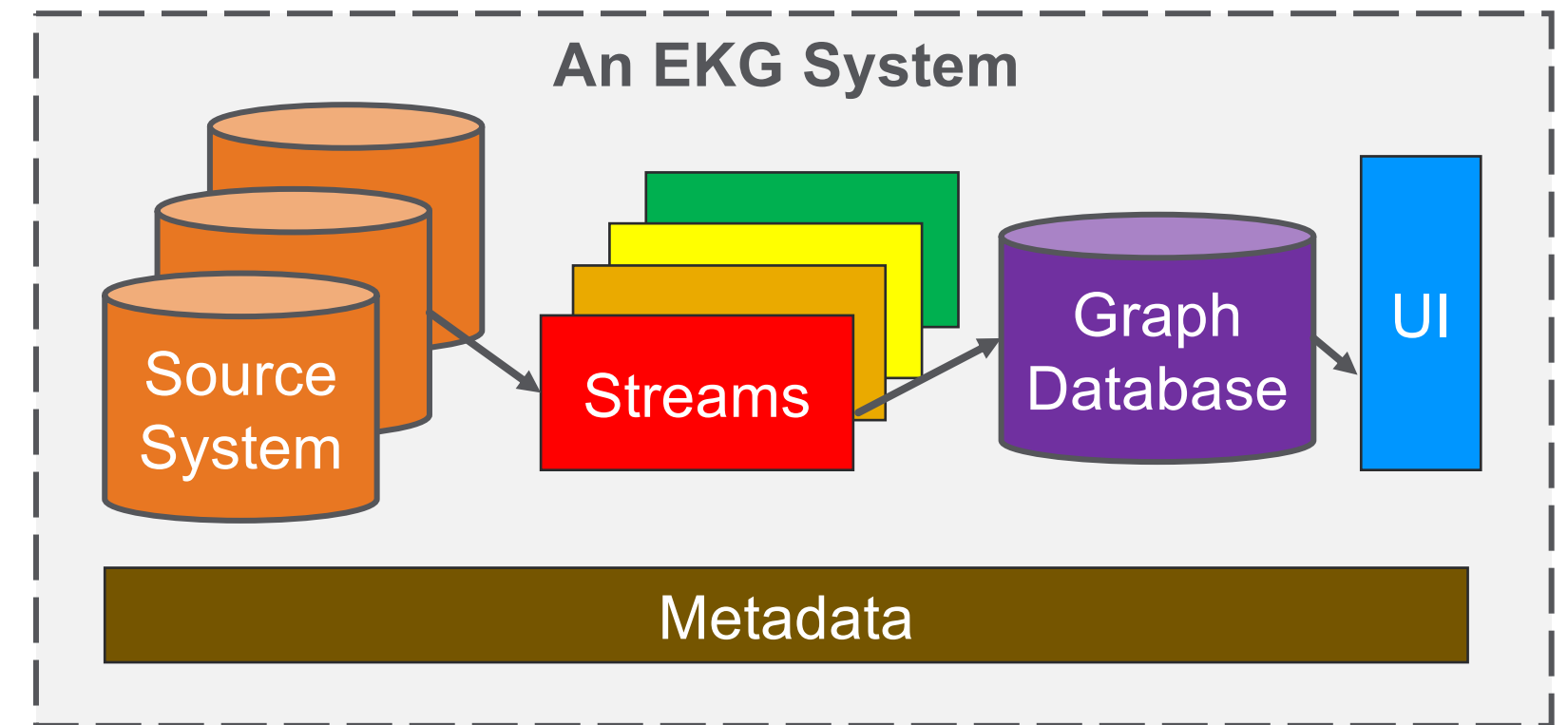
1. right **information**
2. to the right **people**
3. at the right **time (seconds)**

Intelligent triggers

Proactive alerts

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 evolve over **time**
- **Emergent behavior** arises in complex systems



What is Systems Thinking?



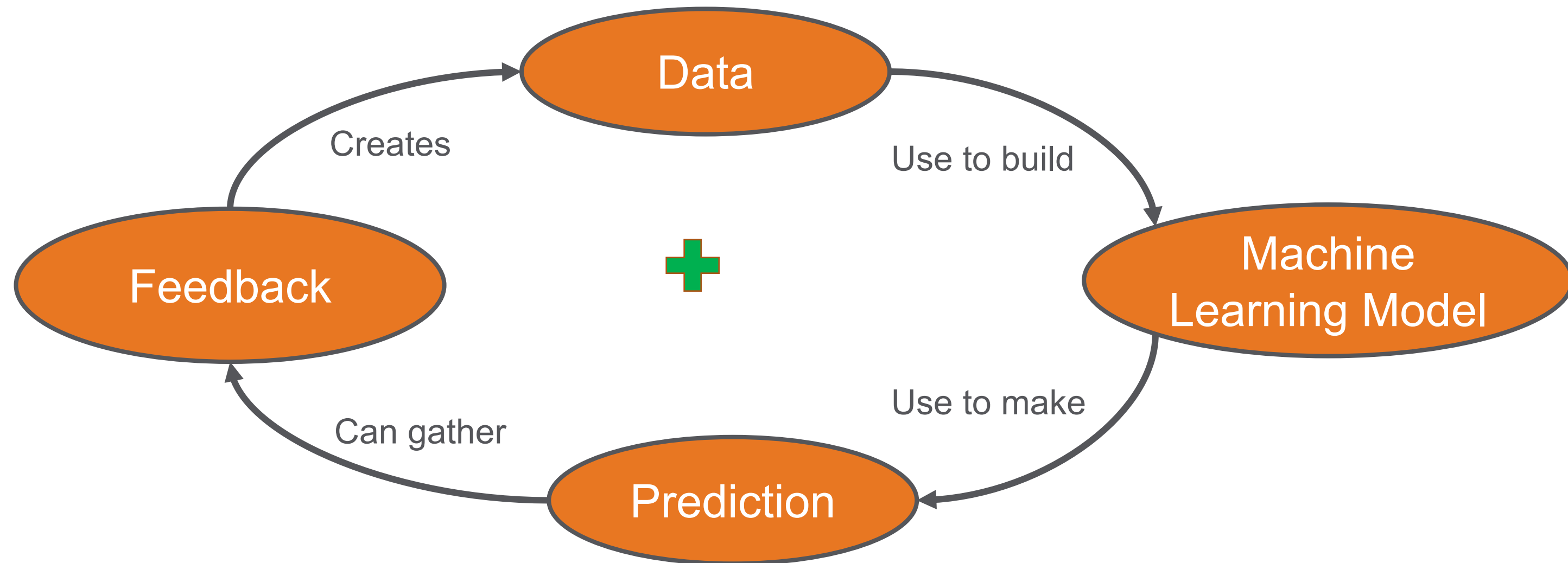
Systems thinking is a holistic **approach to analysis** that focuses on the way that a **system's** constituent parts interrelate and how systems work over **time** and within the context of larger systems.

A Definition of Systems Thinking: A Systems Approach

Ross D. Arnold and Jon P. Wade

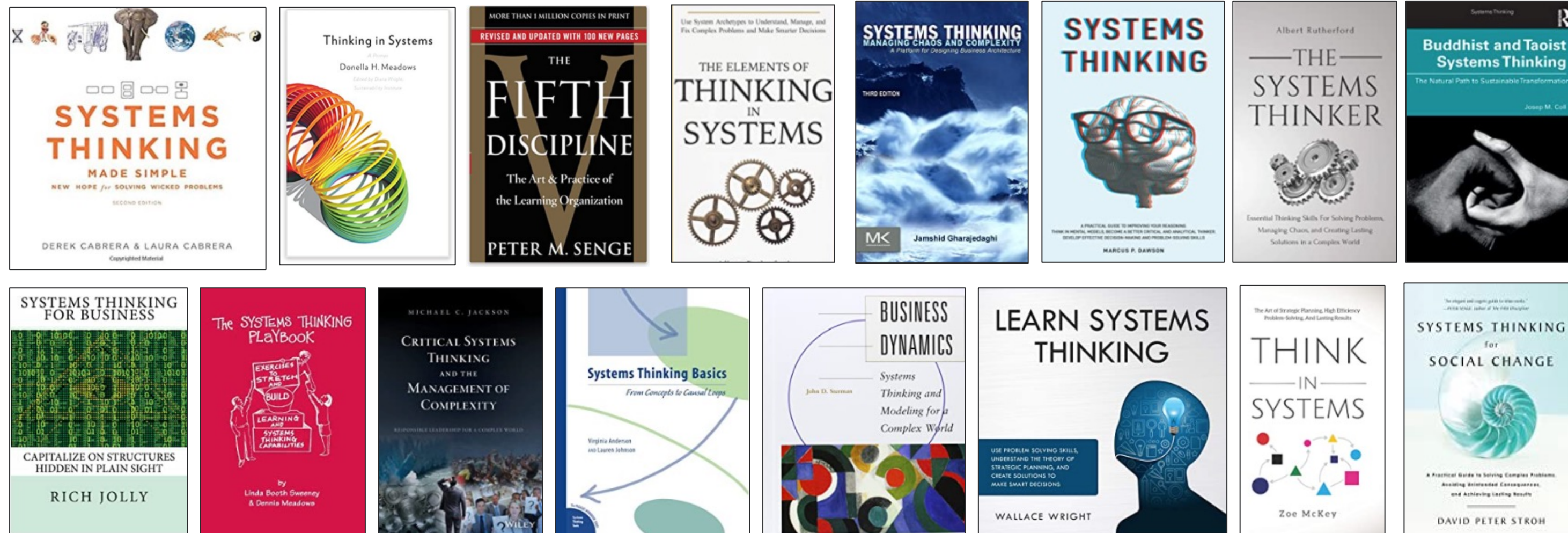
<https://www.sciencedirect.com/science/article/pii/S1877050915002860>

The AI Flywheel



- More data creates more precise machine learning models
- Feedback is central to getting more data
- The trick is getting started!

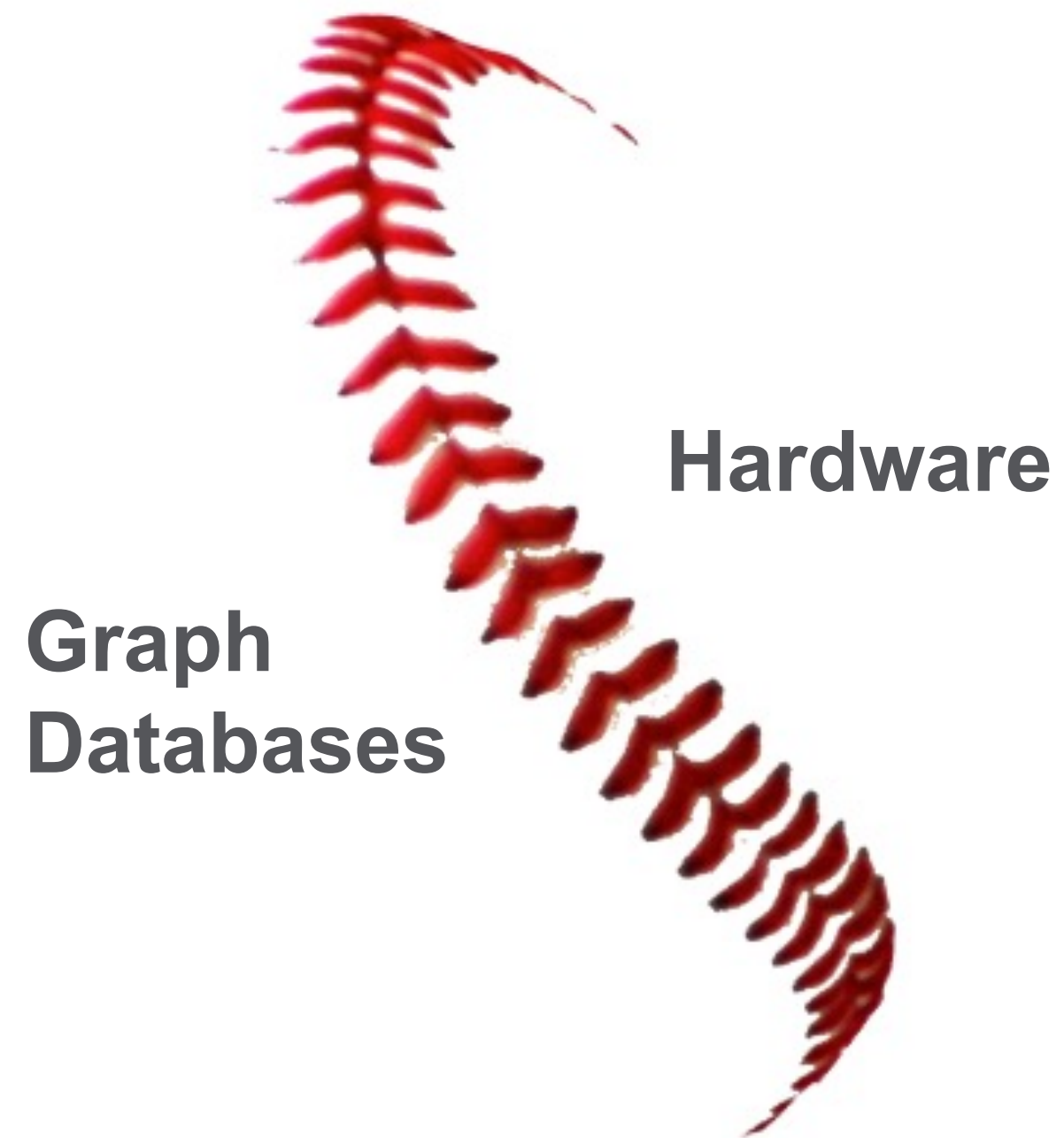
Wonderful Resources for Learning Systems Thinking



- 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

EKG Software and EKG Hardware

A Systems Thinking Example



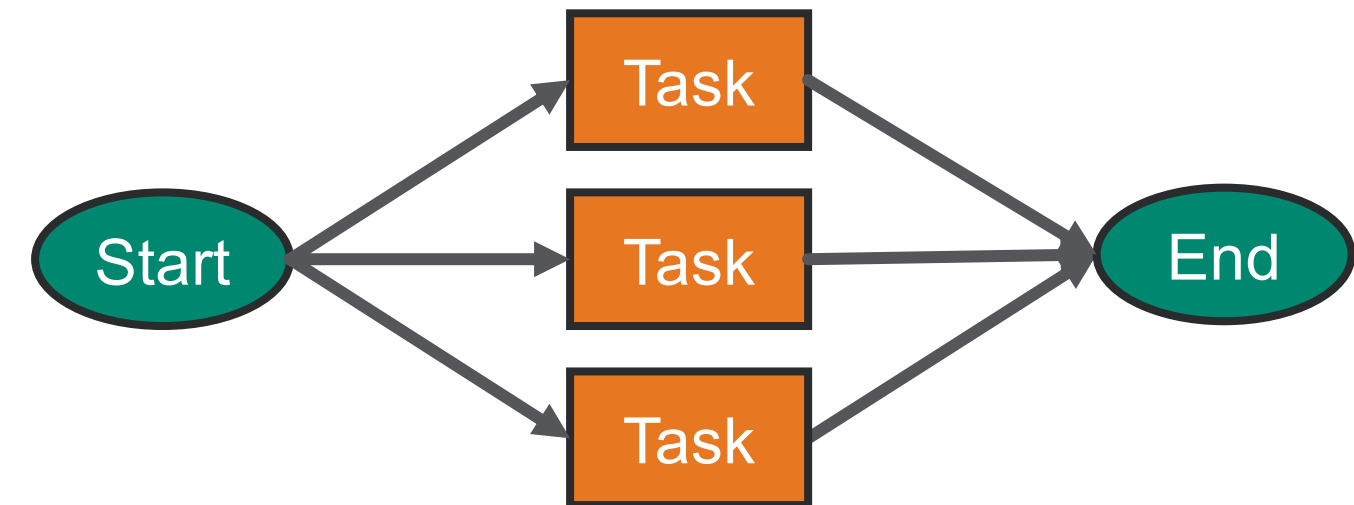
Serial vs. Parallel Graph Algorithms

Serial Graph Algorithms



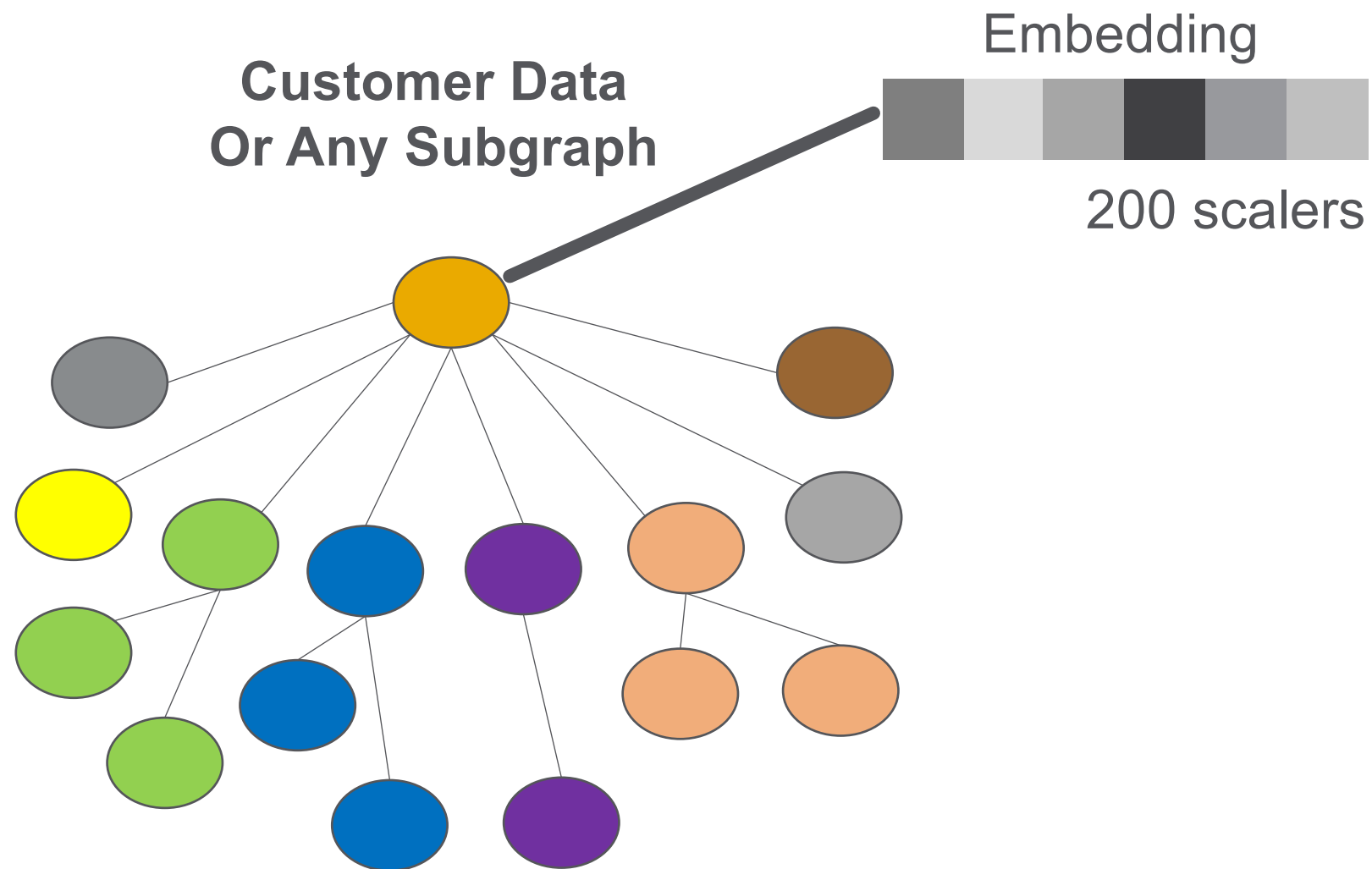
- One task cannot begin before the prior task is complete
- Task **order** is important
- Serial algorithms work well on traditional CPUs

Parallel Graph Algorithms



- Many tasks can be done independently
- Task order is not relevant
- Tasks can usually be done faster on GPU or FPGAs

Graph Embedding

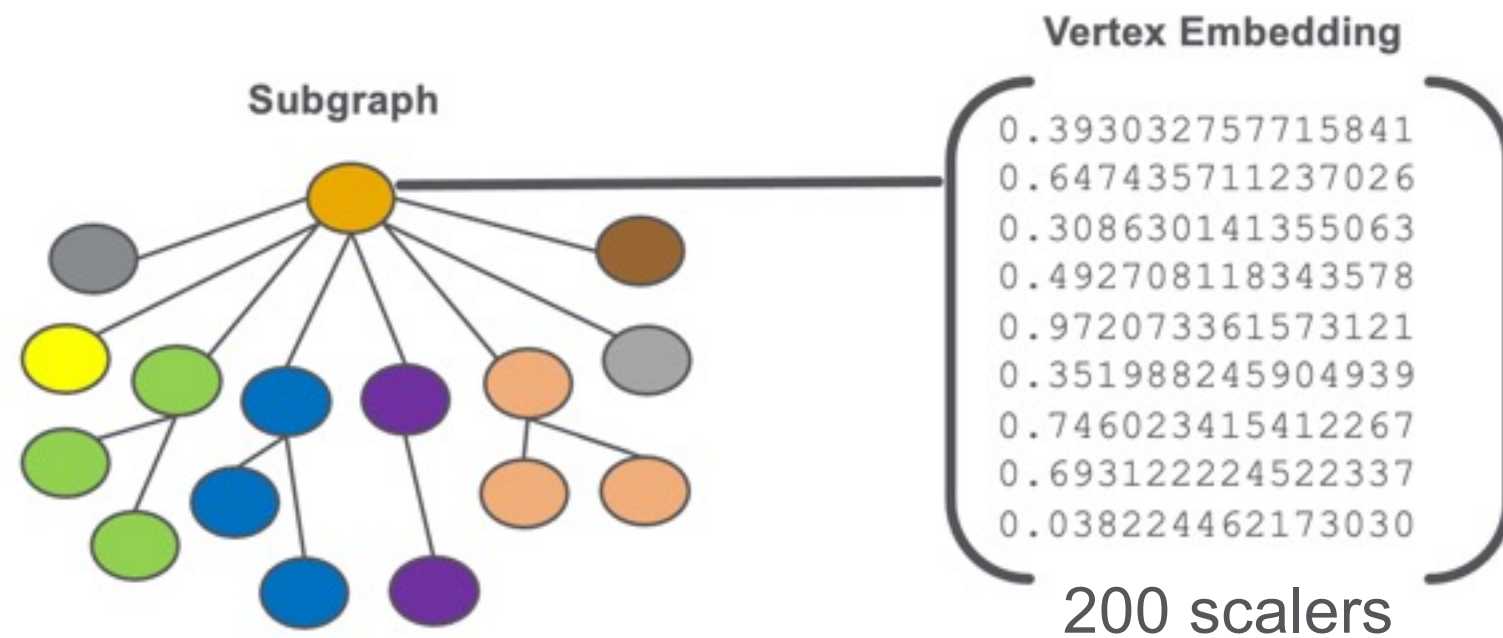


- A data structure used to create fast similarity calculations
- Usually stored as a fixed-length vector of scaler decimals
- Optimized for fast parallel comparison

Understanding Graph Embeddings:

<https://dmccreary.medium.com/understanding-graph-embeddings-79342921a97>

Vertex Embedding

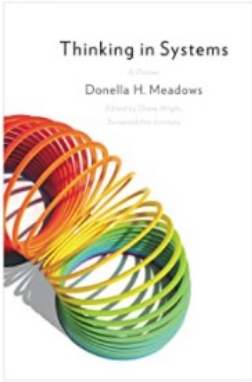


- An embedding is a vector of scalar numbers used to help you find similarity items
- Graph embeddings consider not just the attributes of a vertex, but the **context** of the vertex in a graph
- Embeddings are learned representations of the knowledge in an enterprise knowledge graph (EKG)
- The larger the graph, the more precise the embedding encodes information about the role of that vertex in the community of all other vertices of similar type
- Machine learning is used to create embeddings within EKGs
- You can "do math" on customers
 - add, subtract
 - average

Customer Experience Depends on Real Time Recommendations

Your item:

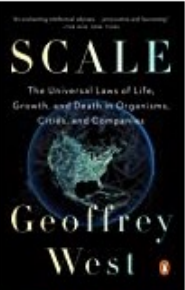
Best Seller



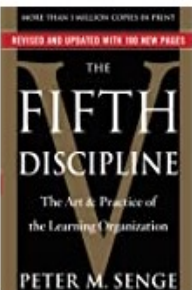
Thinking in Systems: A Primer
by Donella H. Meadows and Diana Wright
★★★★☆ ~ 1,738
Paperback
\$13⁰⁰ \$19.95

Real time similarity calculations

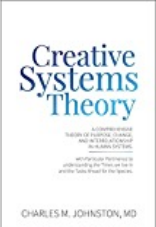
Products **related** to this item



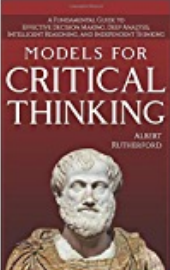
Scale: The Universal Laws of Life, Growth, and Death in...
>Geoffrey West
★★★★☆ 549
Paperback
\$9.99



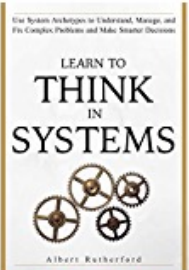
The Fifth Discipline: The Art & Practice of The Learning Organization
>Peter M. Senge
★★★★☆ 1,513
Paperback
\$20.46



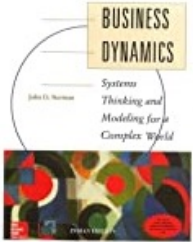
Creative Systems Theory: A Comprehensive Theory of Purpose, Change, and Interrelati...
Charles Johnston
The definitive guide to how and why human systems evolve.
★★★★☆ 14
Kindle Edition
\$4.95



Models For Critical Thinking: A Fundamental Guide to Effective Decision Making, Deep...
Albert Rutherford
★★★★☆ 154
Paperback
\$16.99 ✓prime

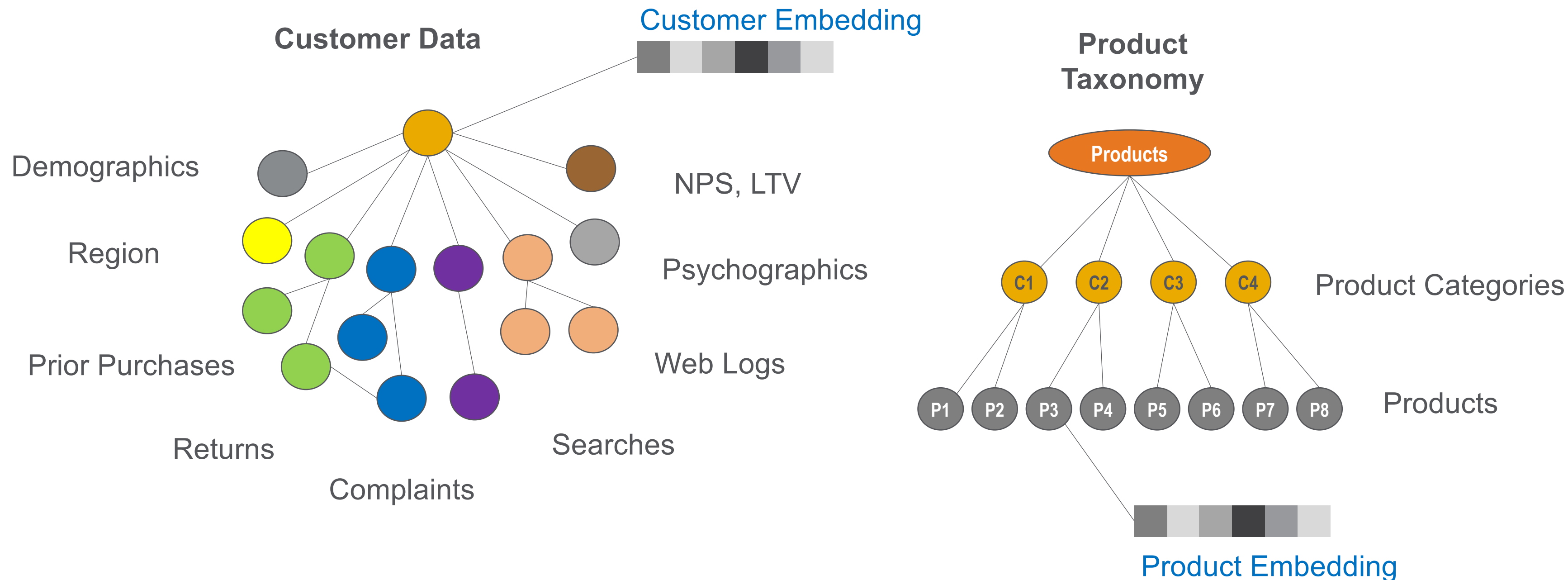


Learn To Think in Systems: Use System Archetypes to Understand, Manage,...
Albert Rutherford
★★★★☆ 186
Kindle Edition
\$5.99



Business Dynamics
STERMAN
★★★★☆ 47
Paperback
\$28.99

Product Recommendation

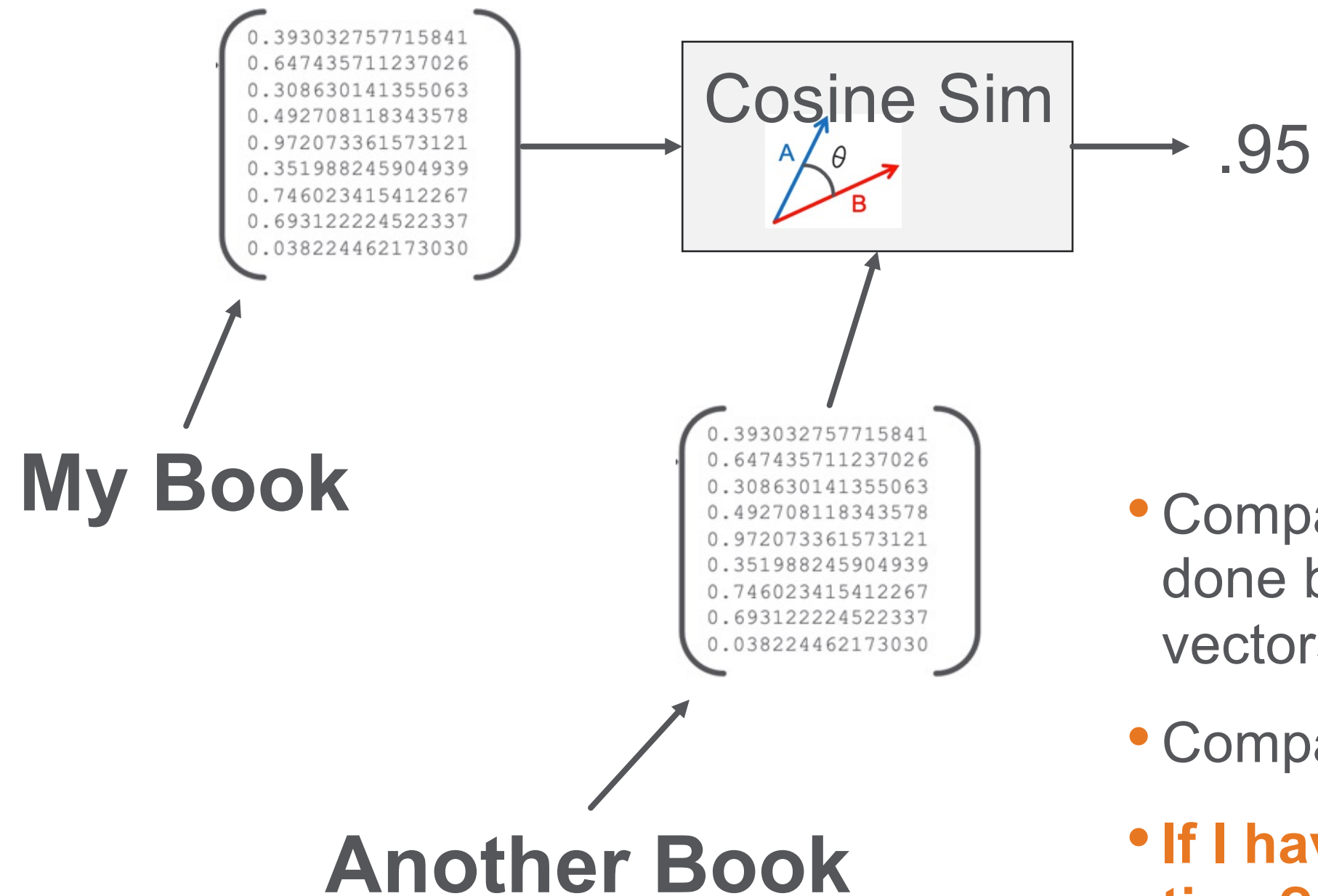


Given a **context**, what similar customers buy what similar products?

NPS – Net promoter score

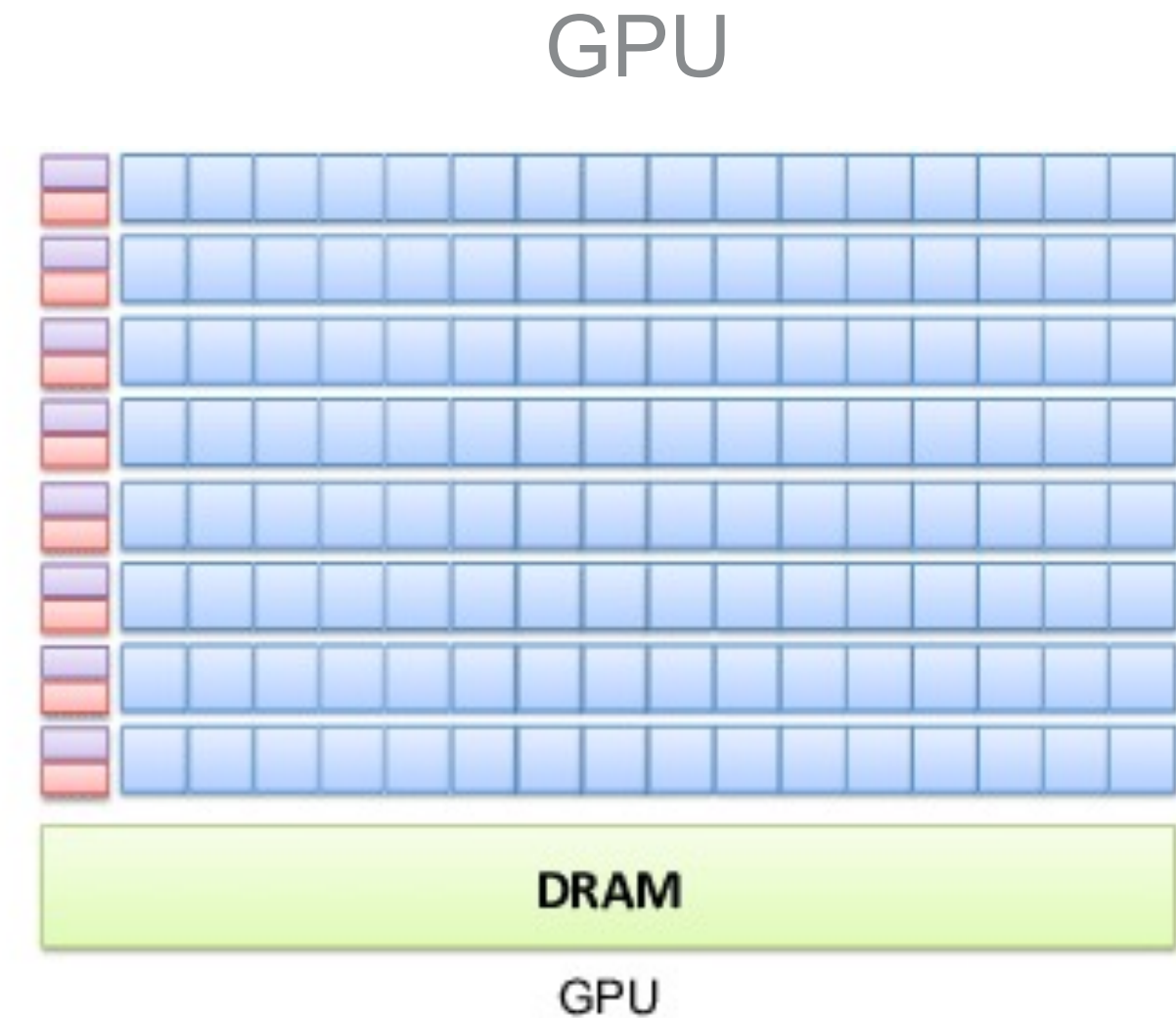
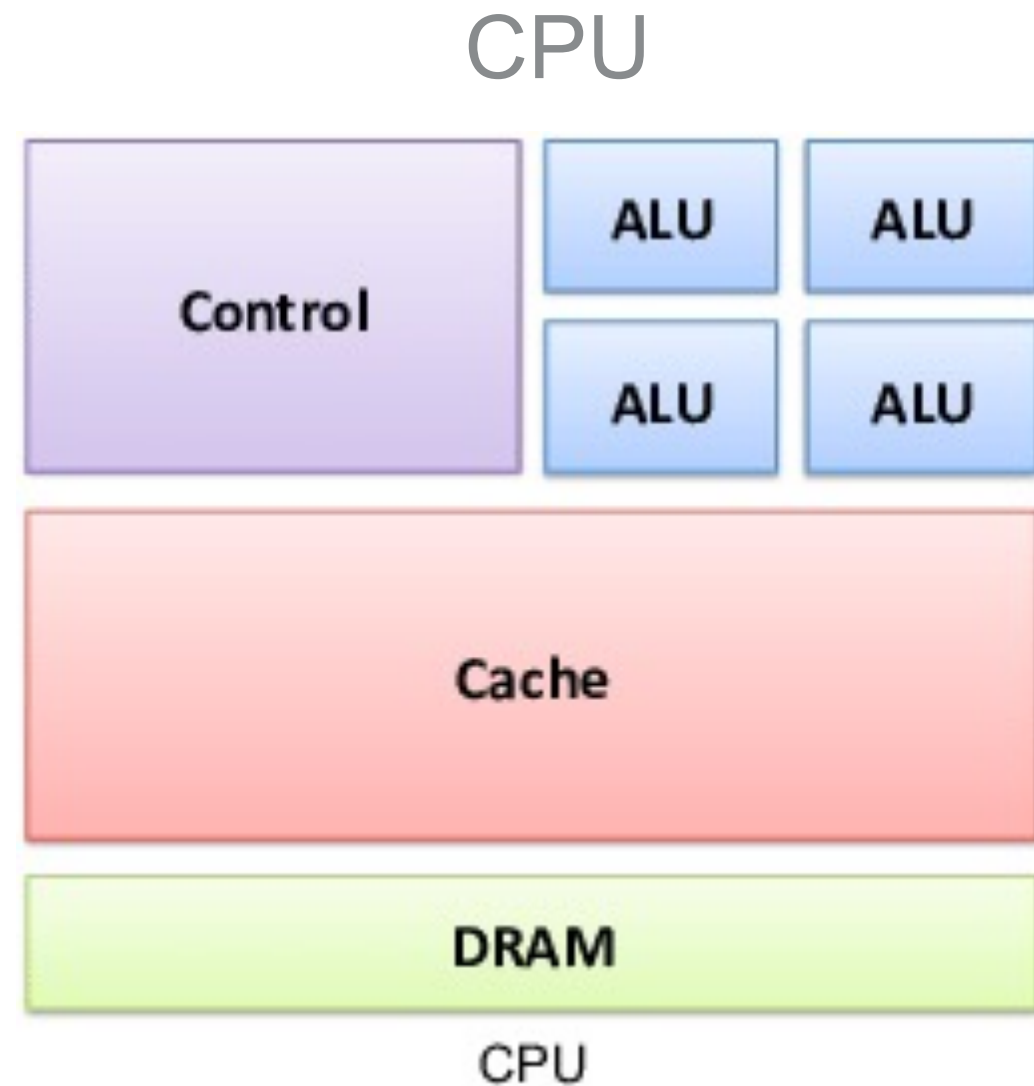
LTV - Lifetime value of a customer

Cosine Similarity – My Favorite Graph Algorithm



- Comparison between any two items can be now be done by similarity comparison between the item vectors
- Comparisons can be done in parallel!
- **If I have 10 million books can we do this is real-time?**

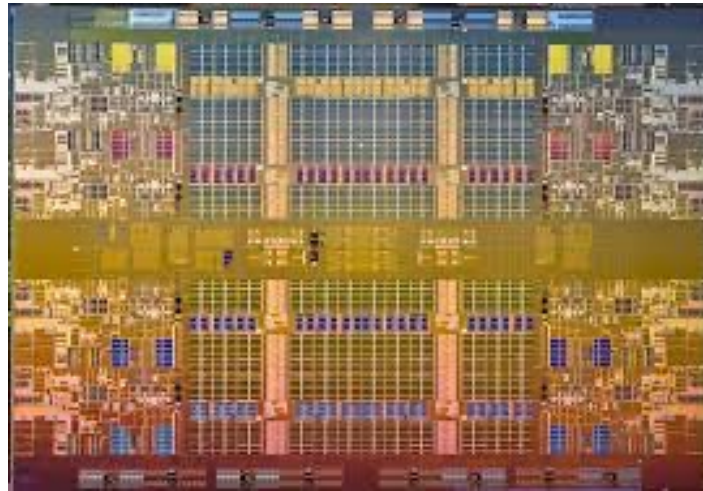
GPUs and SIMD are Ideal for “Euclidian” Geometry Problems



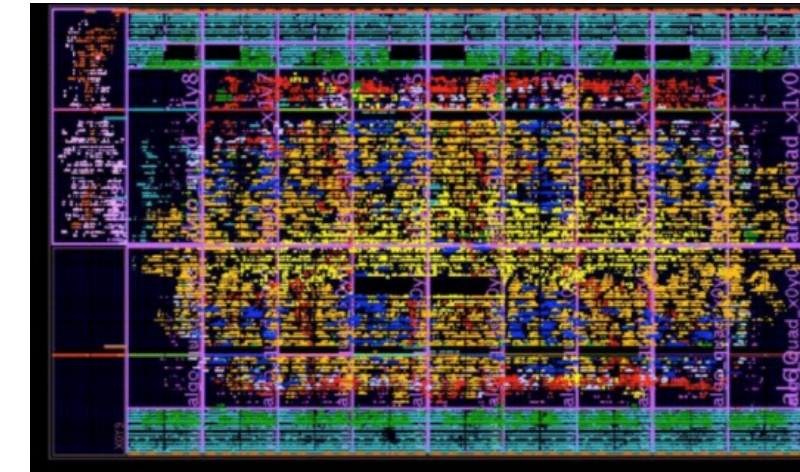
- Large number of instructions
- X86 has over 1,500 instructions

- GPUs have a single instruction that manipulates uniform data (SIMD)
- All threads execute the exact same instruction

The Field Programmable Gate Array (FPGA)?



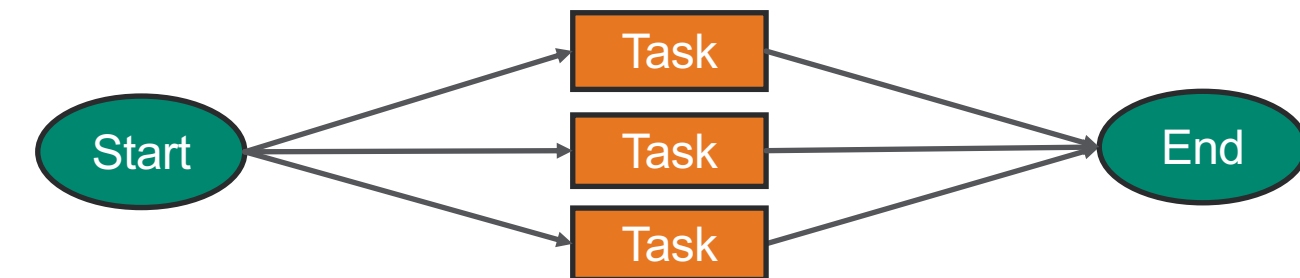
Intel Xeon



Xilinx FPGA



- 50 billion transistors that have **fixed** wiring
- Fixed floorplan design
- Ideal for **serial** processing



- 50 billion transistors we can **rewire** in 2 seconds
- Reconfigurable floorplan
- Ideal for **parallel** processing
- Cosine similarity on 10M items in under 50 msec

The “BERT BOX” Vendors

The Incumbent:



IPOD-16 created BERT for English Wikipedia in 12 hours

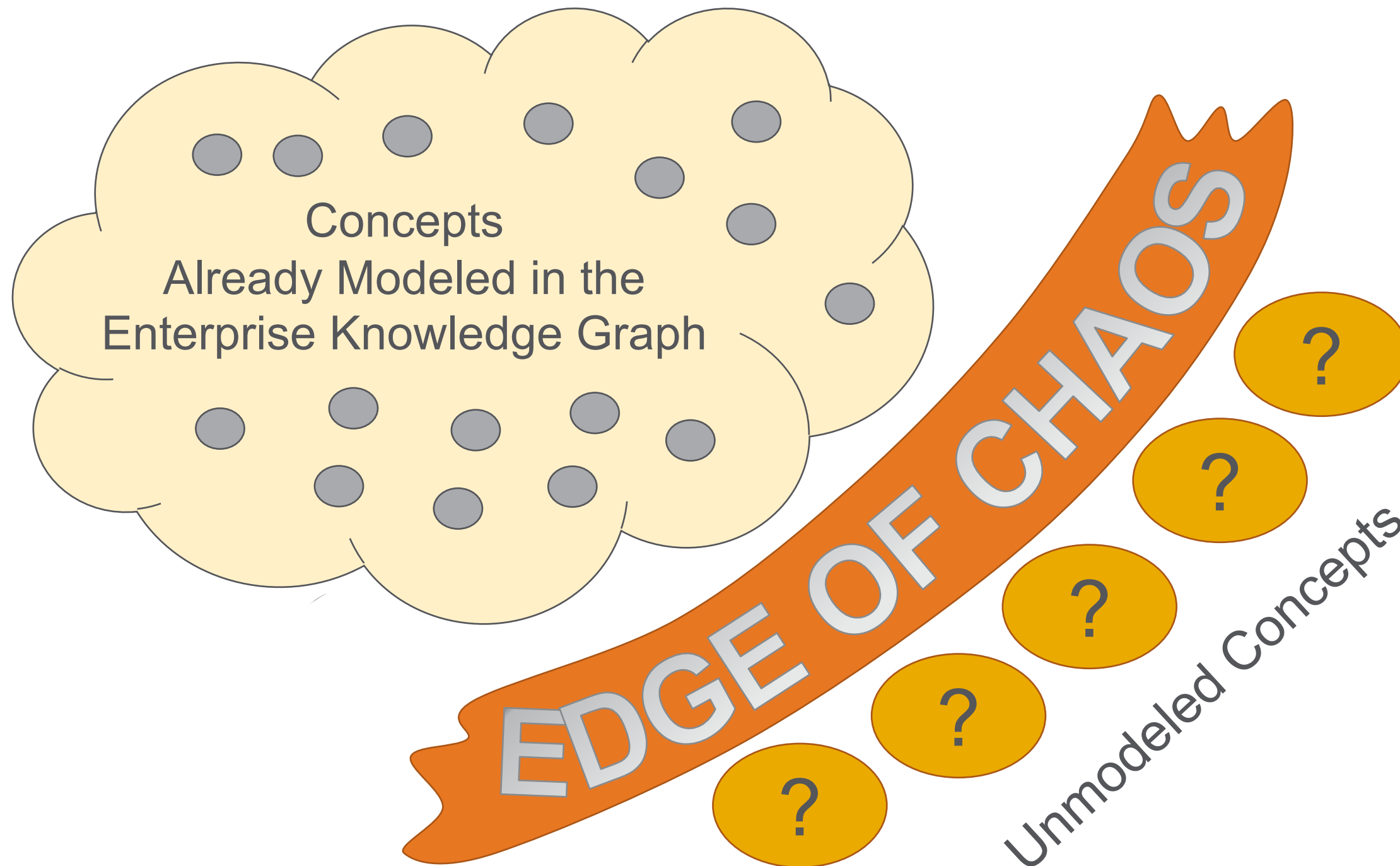
The Challengers:



\$1.3 Billion in funding for new AI Hardware in 2021



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.

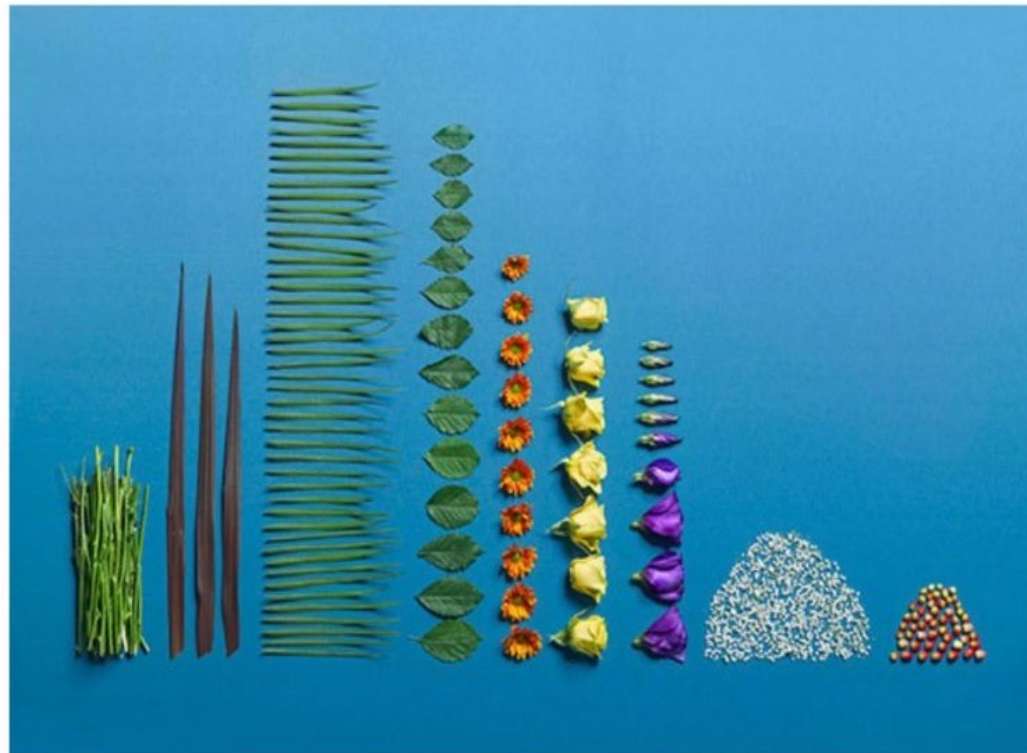
2) The part of the world that you have not modeled yet. We call this the “region of chaos”.

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

Conclusion

Siloed Thinking



Systems Thinking



- Systems Thinking helps us get out of “siloed” locally optimized thinking
- We need to take a holistic view of our EKG strategies to help drive incremental value
- Systems Thinking can help us get faster paths to insights and value
- We need to continue to build **stories** about how Systems Thinking helps us use EKGs to help us serve our customers

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://dmccreary.medium.com/>



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



@dmccreary

