










Unraveling the Hidden Layers: Graphs and Machine Learning for Workforce Planning

Andy Korinda



Agenda

-  About HSC & me
-  Headcount & expansions
-  Path to inspiration
-  Data science lite
-  Graphs aren't just pictures
-  Guts of machine learning
-  Results & To come

What do we do?

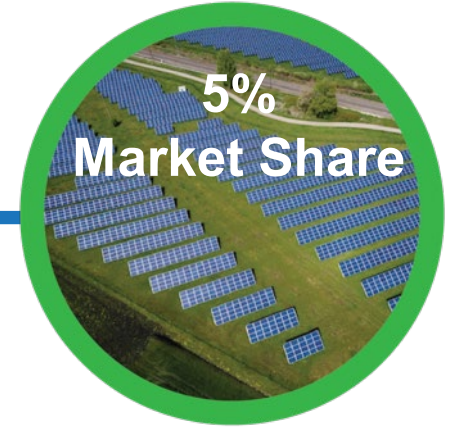
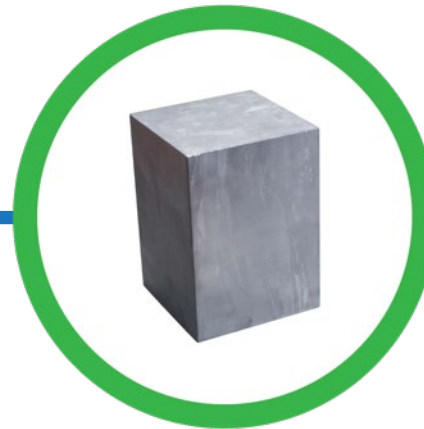


Polysilicon: Foundation of Our High-Tech Future

The electronics supply chain



The solar supply chain





1,350
workers on site

Chemical & Mechanical
Process

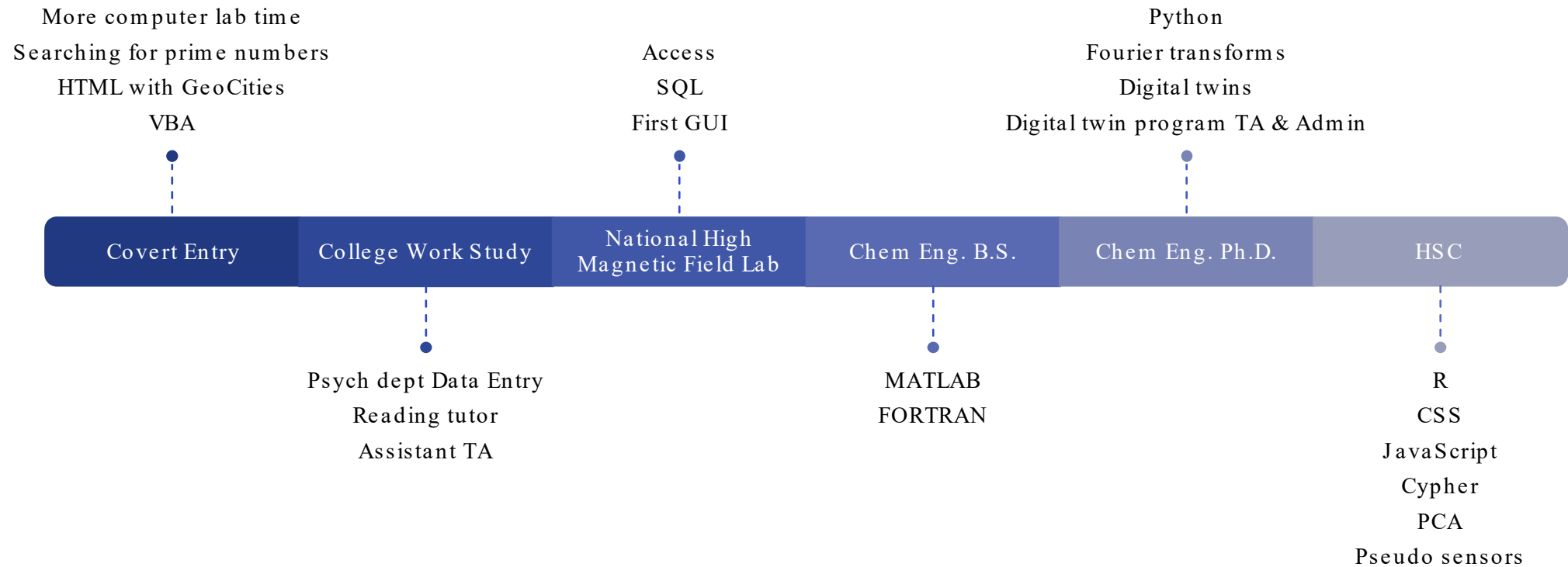
Manual & Automated
Equipment

Earth's *purest* manmade
material
99.999999999%



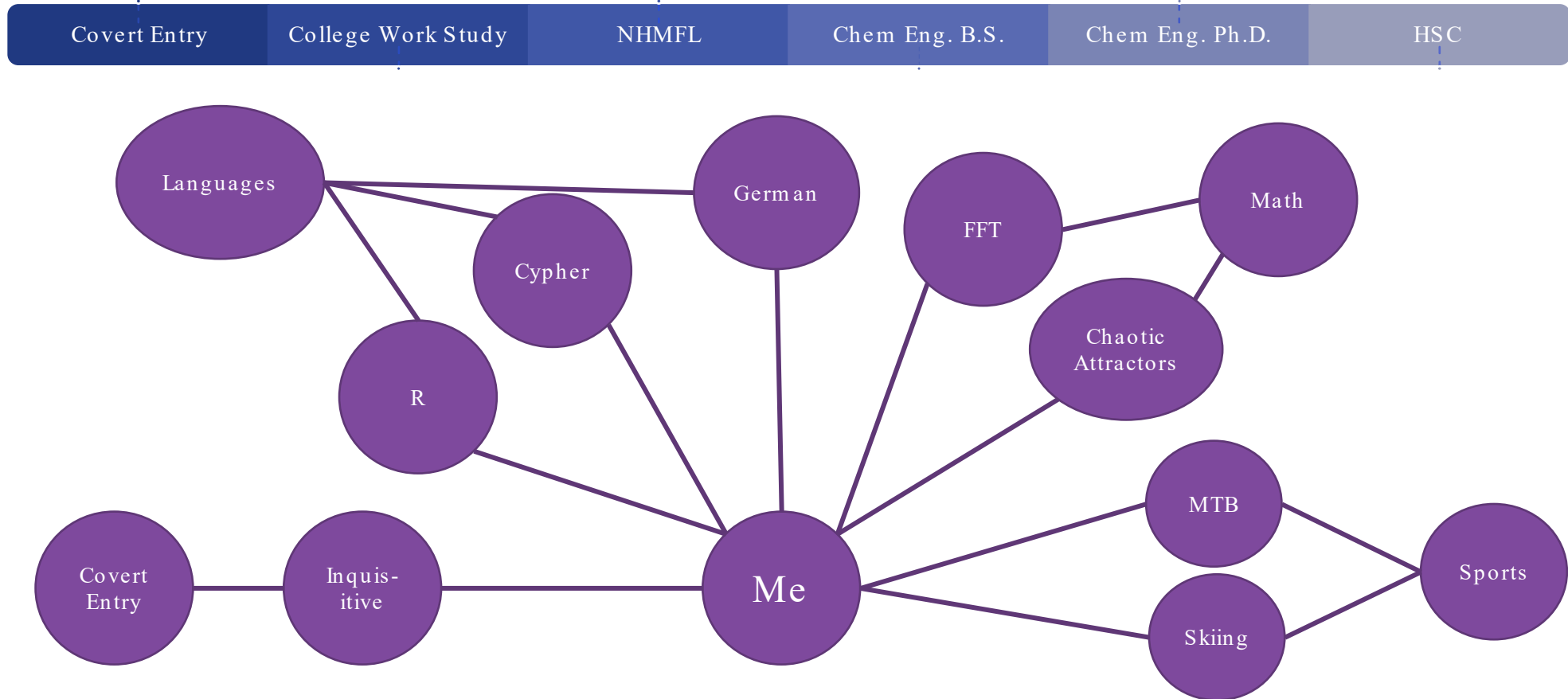
WHO AM I?

A timeline tells a story



WHO AM I?

A graph presents connections



PROBLEM OF TIMELY BACKFILLS

Expansion

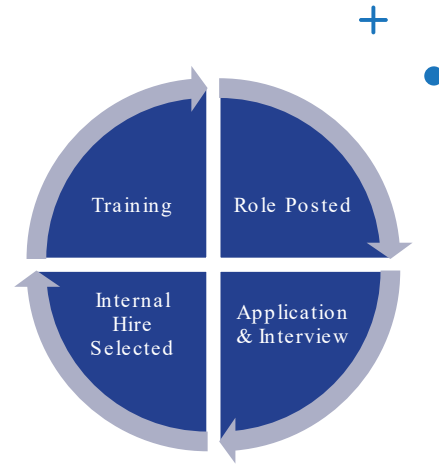
- Preference to promote into the new roles
- 170 new jobs
- 12% headcount increase
- \$375 million investment



Chicken & the Egg

- New role posted
- 1-2 month application & interview process
- Internal hire selected
- 4 to 12 weeks of training to start the new role
- Posting for replacement AFTER training completed
- Repeat process 4+ times to reach entry level positions

*All timelines and job titles are hypothetical



PROBLEM OF TIMELY BACKFILLS

Problem

- Up to 20-month process to complete the backfill chain
- Plus: Replacement of normal turnover of existing roles
- Orders already received based on expansion capacity

Solution

- New roles & headcounts proposed
- Project the resulting backfill movements
- Approve the new role, backfill, and turnover postings all at once
- Project training timings & adjust to fit trainer availability/capacity

PATH TO INSPIRATION

Matt Parker “Can you find: five five-letter words with twenty-five unique letters?”

Stand-up Maths YouTube

Code optimization with graphs

Other HSC HR Project

‘23 WiT GLDAS meeting Neo4j

Predicting hiring backfills

“Someone improved my code by 40,832,277,770%”

Meet the team & listen to what is said outside that project.

Conference brainstorming with experts

Graphs
Ontology
Communication

PATH TO INSPIRATION

Matt Parker “Can you find: five five-letter words with twenty-five unique letters?”

Meet the team & listen to what is said outside that project.

Graphs
Ontology
Communication

Stand-up Maths YouTube

Code optimization with graphs

Other HSC HR Project

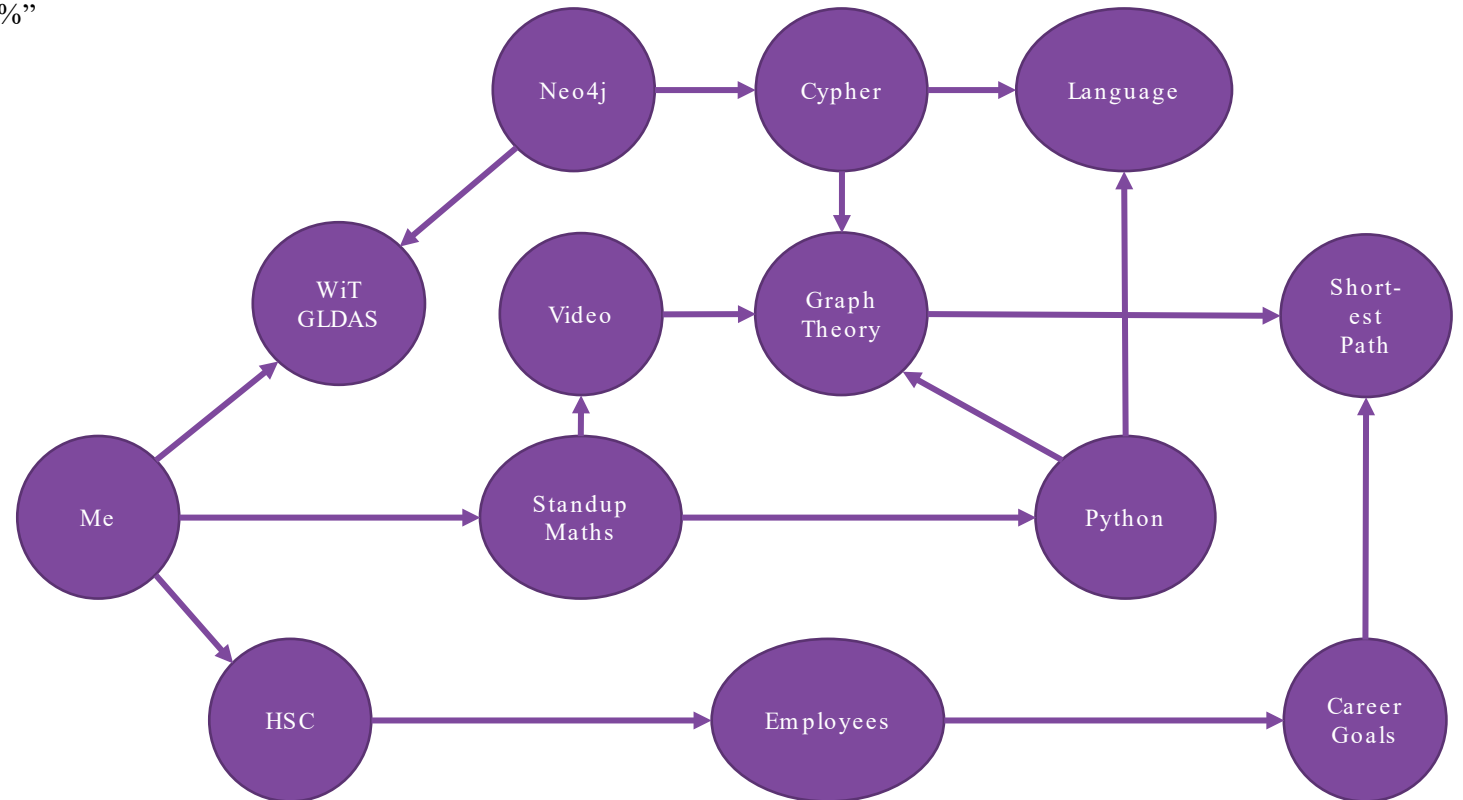
'23 WiT GLDAS meeting Neo4j

Predicting hiring backfills

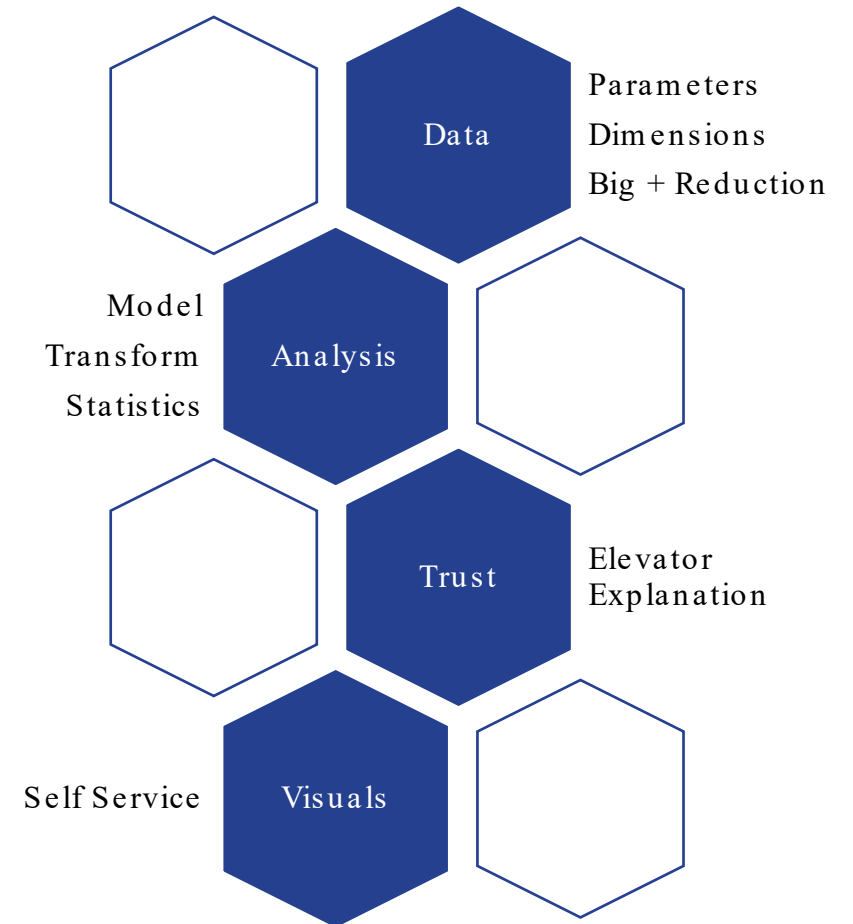
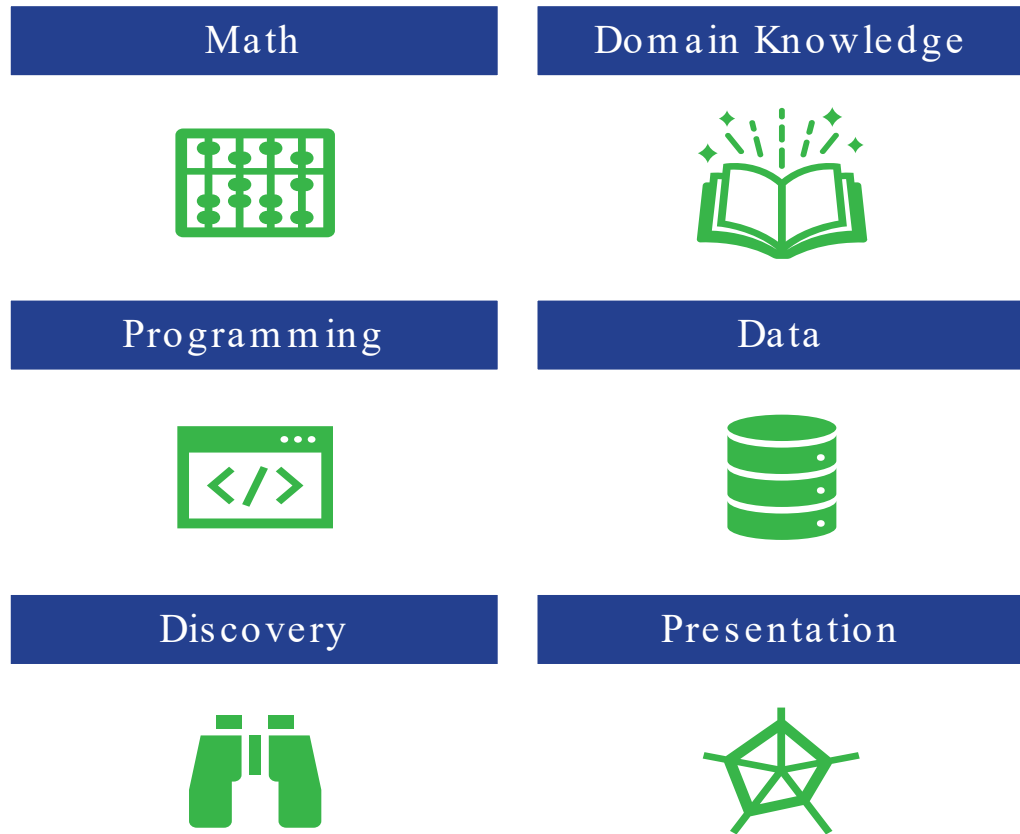
“Someone improved my code by 40,832,277,770%”

Conference brainstorming with experts

- Graph captures more relations than a linear story
- Graph enables alternative queries & representations
- Traversing a graph has speed advantages



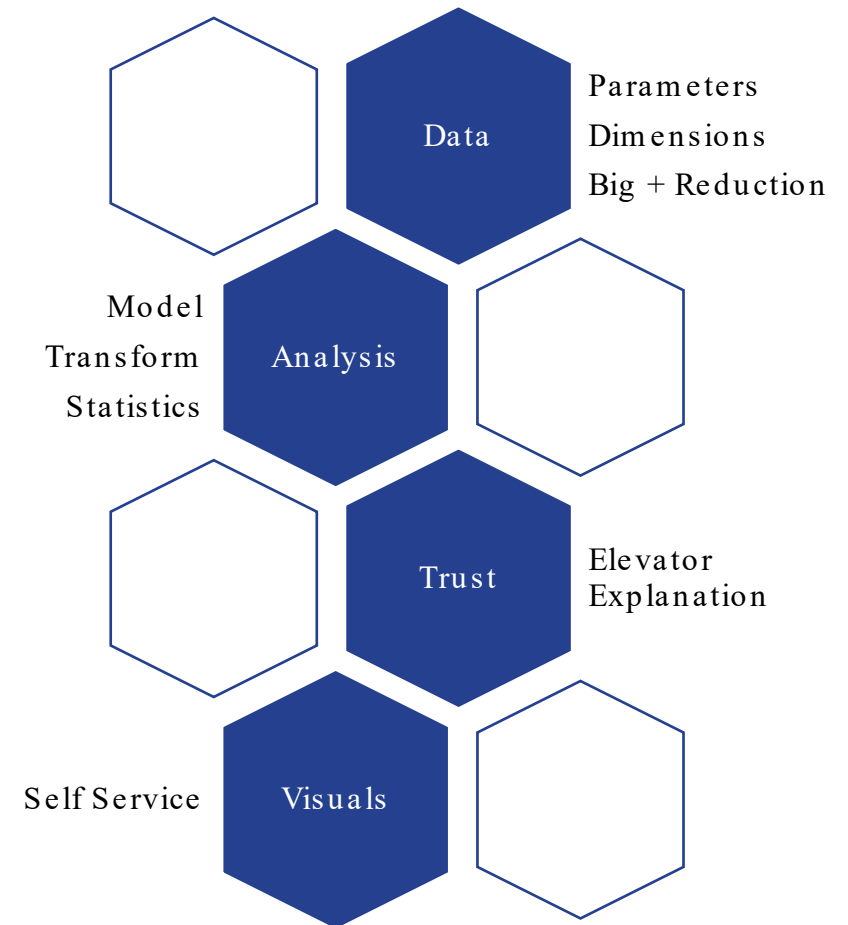
DATA SCIENCE LITE



DATA SCIENCE LITE

Use Data to Find and Evaluate parameter Relationships

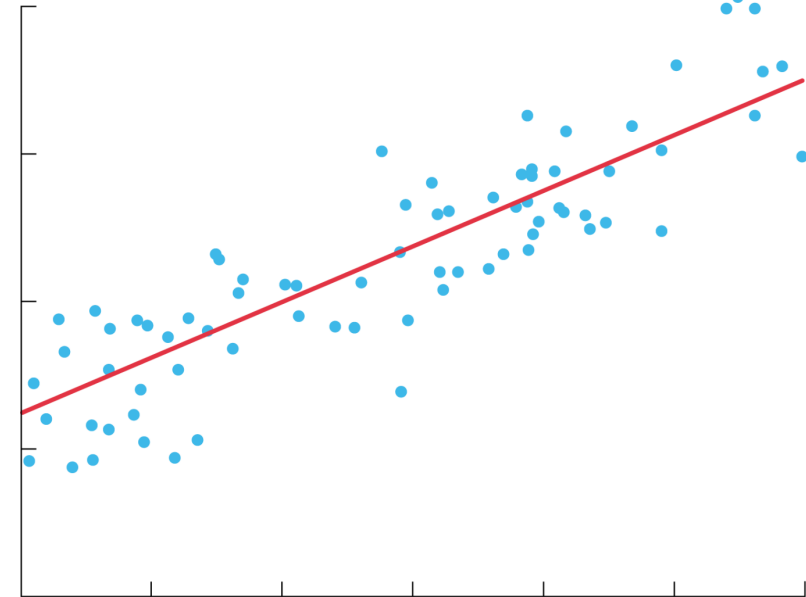
- Find/Model
 - Regression
 - Classification
 - Clustering
- Evaluate
 - Characteristics (statistics)
 - Probability/Confidence/Predictive descriptors
 - Visualize
- Solve Problem / Resolve Hypothesis



DATA SCIENCE LITE

Use Data to Find and Evaluate parameter Relationships

- Find/Model
 - Regression
 - Classification
 - Clustering
- Evaluate
 - Characteristics (statistics)
 - Probability/Confidence/Predictive descriptors
 - Visualize
- Solve Problem / Resolve Hypothesis

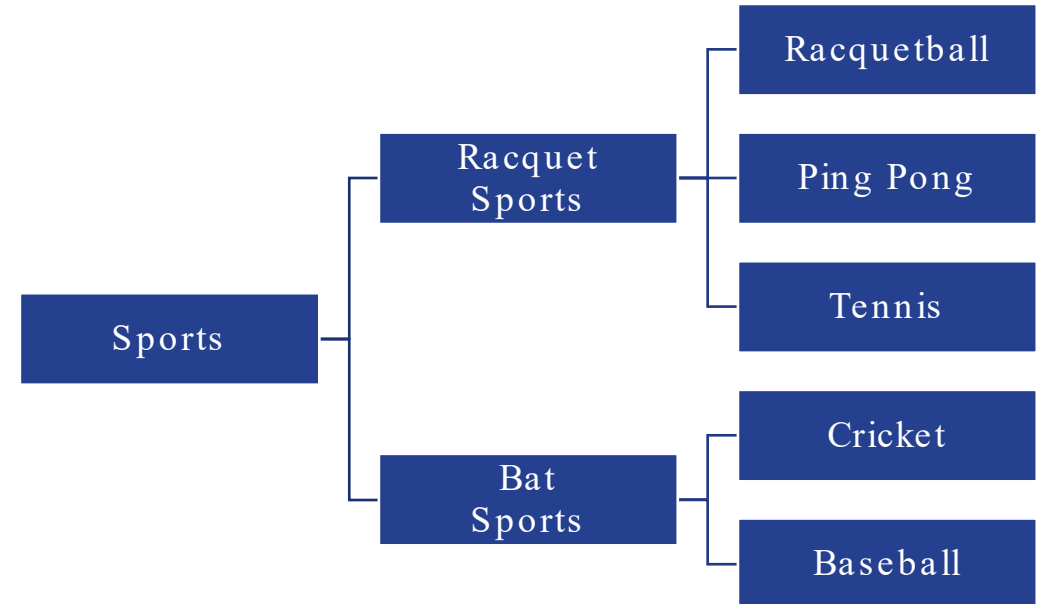




DATA SCIENCE LITE

Use Data to Find and Evaluate parameter Relationships

- Find/Model
 - Regression
 - Classification
 - Clustering
- Evaluate
 - Characteristics (statistics)
 - Probability/Confidence/Predictive descriptors
 - Visualize
- Solve Problem / Resolve Hypothesis





DATA SCIENCE LITE

Use Data to Find and Evaluate parameter Relationships

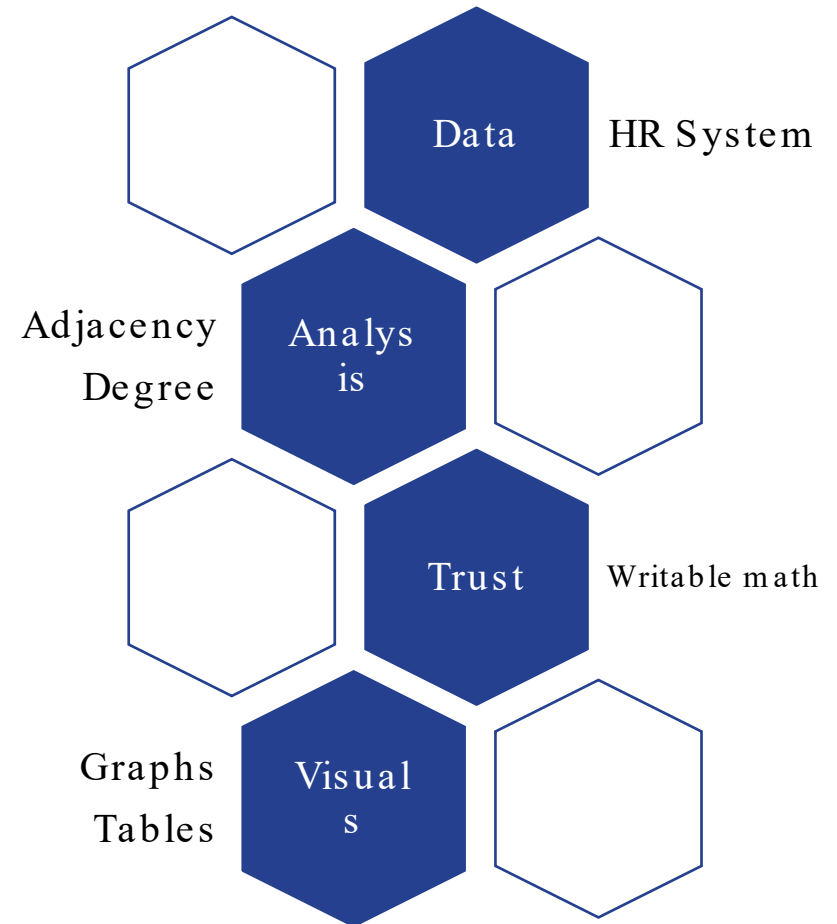
- Find/Model
 - Regression
 - Classification
 - Clustering
- Evaluate
 - Characteristics (statistics)
 - Probability/Confidence/Predictive descriptors
 - Visualize
- Solve Problem / Resolve Hypothesis



DATA SCIENCE For This Talk

Use Data to Find and Evaluate parameter Relationships

- Find/Model
 - Regression
 - Classification
 - Clustering
- Evaluate
 - Characteristics (statistics)
 - Probability/Confidence/Predictive descriptors
 - Visualize
- Solve Problem / Resolve Hypothesis



GRAPH THEORY

Structured pairwise relations
between objects

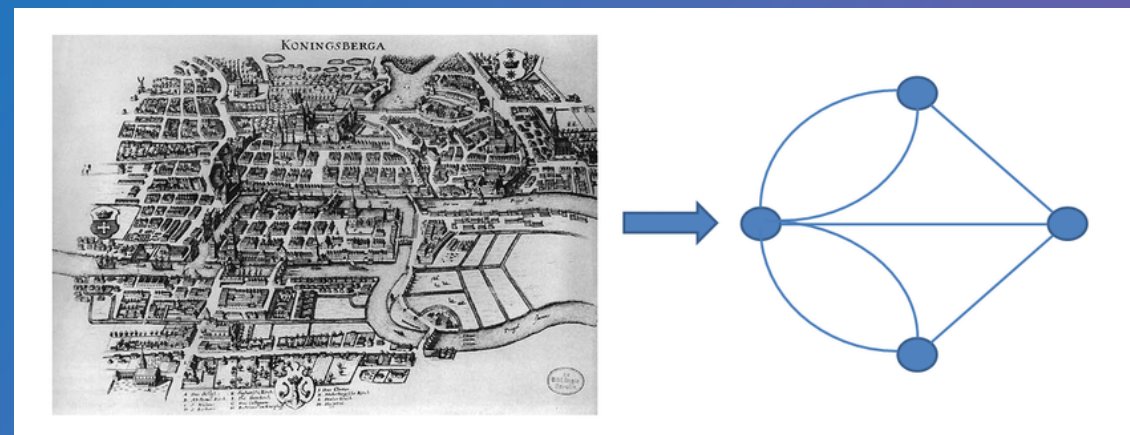
Objects = Nodes = Noun

Relations = Edges = Verb

Nodes and Edges can have attributes (a.k.a. properties)

Directed or Undirected

Ontology is the designed representation of knowledge



Origin: Seven Bridges of Königsberg

School: Social Networks

Today: Algebraic Graph Theory
Application

adds first principal boundaries

GRAPH THEORY

Structured pairwise relations
between objects

Objects = Nodes = Noun

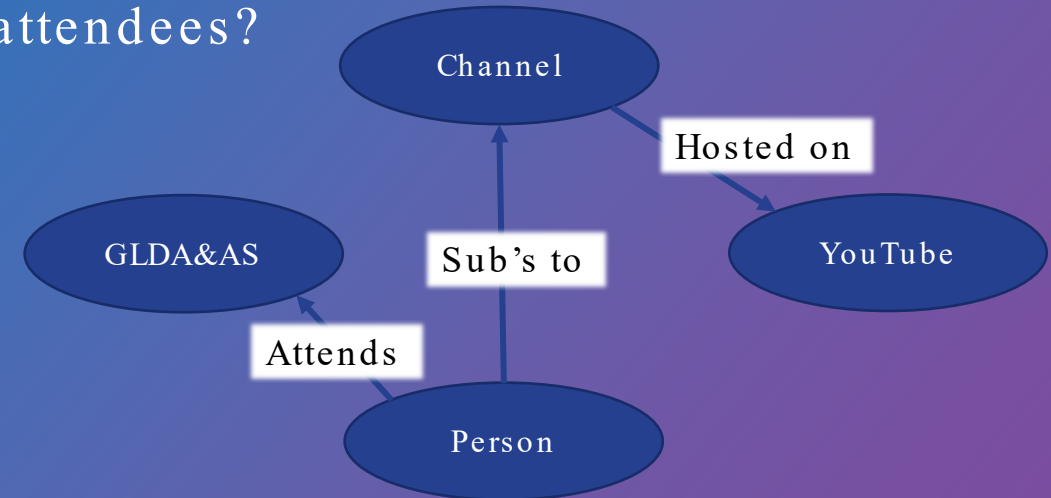
Relations = Edges = Verb

Nodes and Edges can have attributes (a.k.a. properties)

Directed or Undirected

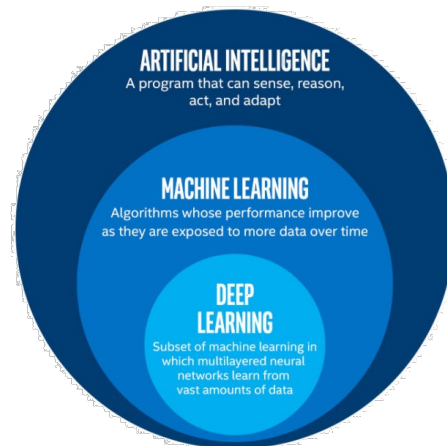
Ontology is the designed representation of knowledge

What is the most
subscribed YouTube
channel by today's
attendees?



GUTS OF MACHINE LEARNING

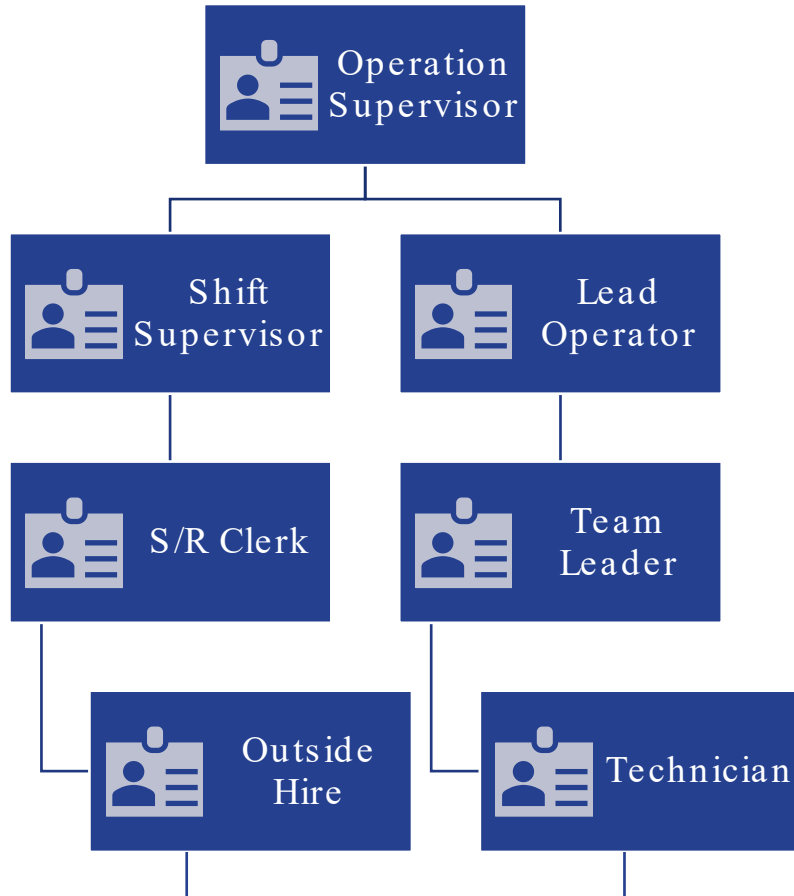
- ML layers are matrices
- Layers are connected by linear algebra
- With career paths we see the guts of AI/ML/NN



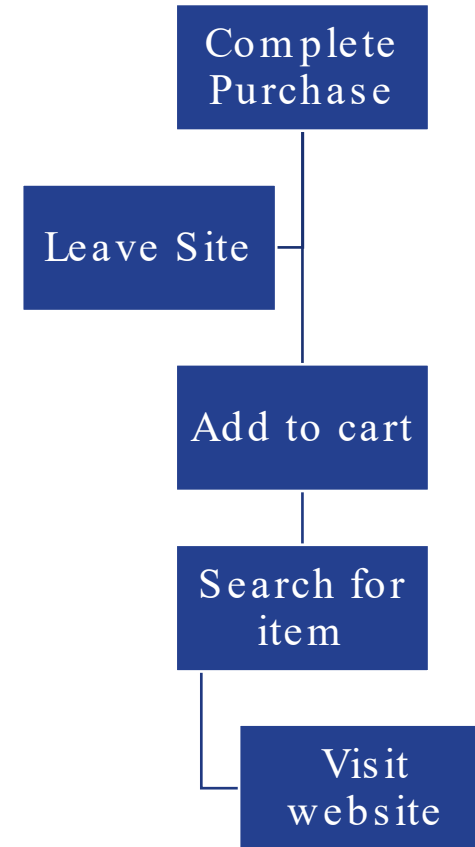
“The pile gets soaked with data and get mushy over time, so technically it’s recurrent.”

A PATH IS A GRAPH

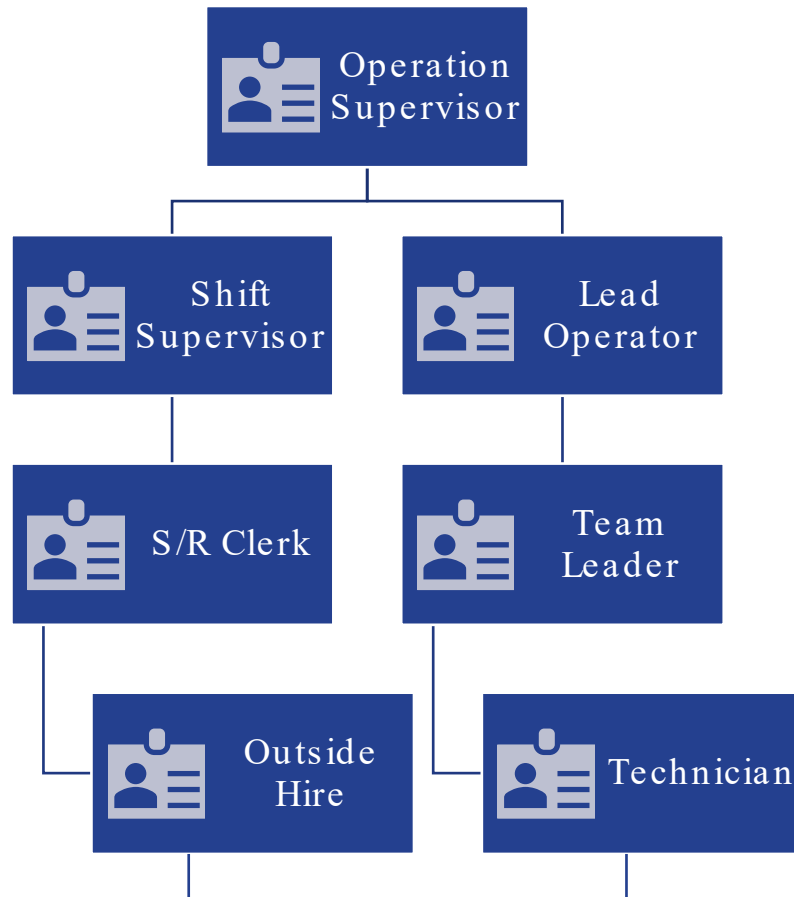
Careers



Web Purchase



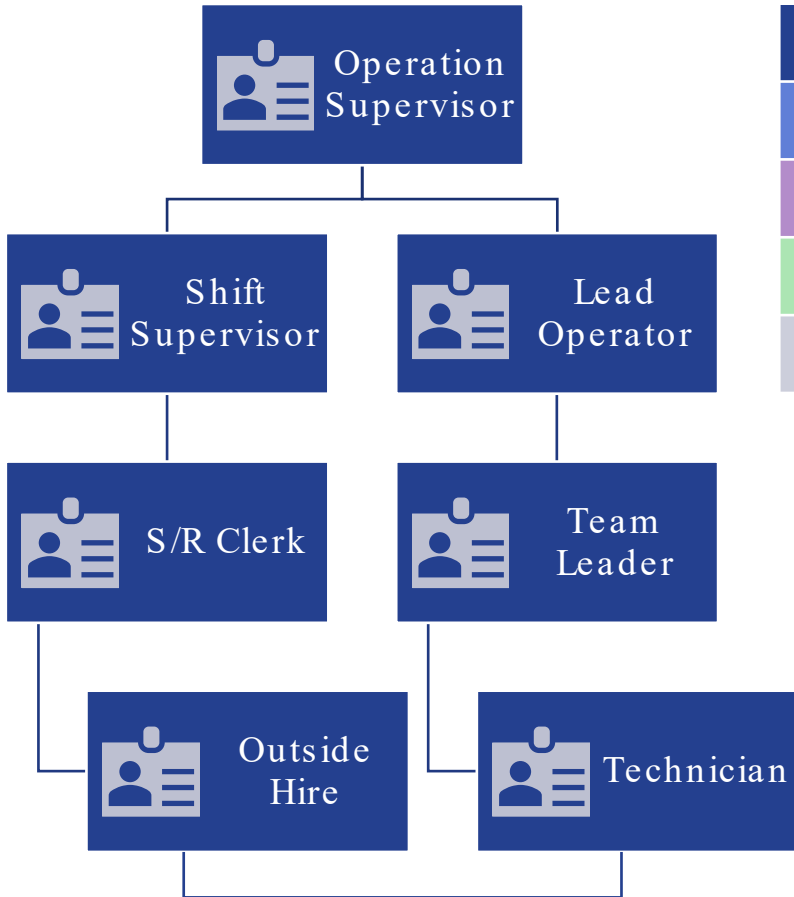
A GRAPH IS A MATRIX



OS						OS -> OH
SS -> OS	SS			SS -> TL		SS -> OH
LO -> OS		LO	LO -> S/R			LO -> OH
	S/R -> SS		S/R			S/R -> OH
		TL -> LO		TL		TL -> OH
				Tech -> TL	Tech	Tech -> OH
			OH -> S/R		OH -> Tech	OH



A MATRIX IS DESCRIPTIVE



- Positions
- Promotions
- Demotions
- Outside Hires
- Separations

OS						OS -> OH
SS -> OS	SS			SS -> TL		SS -> OH
LO -> OS		LO	LO -> S/R			LO -> OH
	S/R -> SS		S/R			S/R -> OH
		TL -> LO		TL		TL -> OH
				Tech -> TL	Tech	Tech -> OH
			OH -> S/R		OH -> Tech	OH

A MATRIX IS ALGEBRAIC

OS						OS -> OH
SS -> OS	SS			SS -> TL		SS -> OH
LO -> OS		LO	LO -> S/R			LO -> OH
	S/R -> SS		S/R			S/R -> OH
		TL -> LO		TL		TL -> OH
				Tech - > TL	Tech	Tech - > OH
			OH -> S/R		OH -> Tech	OH

 \times

[OS Position SS Position LO Position S/R Position TL Position Tech Position Outside Hire]
---	---	---

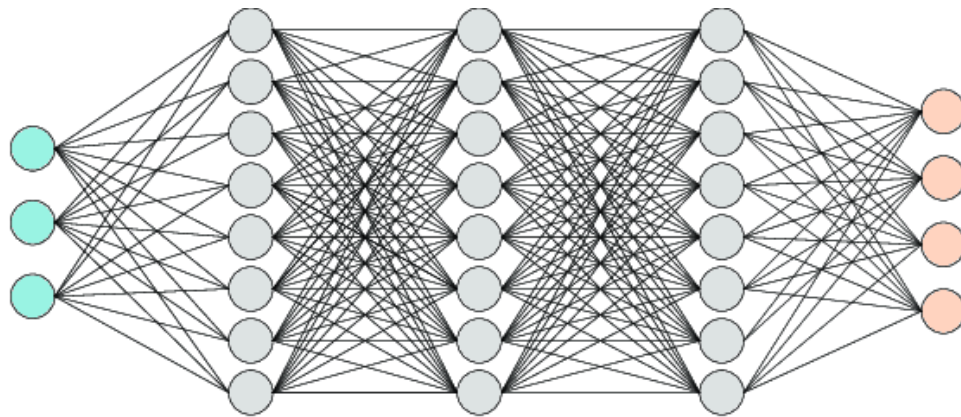
 $=$

[#OS Needed #SS Needed #LO Needed #S/R Needed #TL Needed #Tech Needed Optional]
---	---	---

MATRICES ARE STACKABLE

- Stack to view over time
(Matrix \rightarrow Tensor)
- Mechanically same as:
 - Neural Network layers
 - Generative AI response token progression

$$\begin{array}{c}
 \begin{bmatrix} \text{OS} & \dots & 1 & \text{[OS Position]} & \text{[#OS Needed]} \\ \vdots & & & & \end{bmatrix} \\
 \begin{bmatrix} \text{OS} & \dots & 1 & \text{[OS Position]} & \text{[#OS Needed]} \\ \vdots & & & & \end{bmatrix} \\
 \begin{bmatrix} \text{OS} & \dots & 1 & \text{[OS Position]} & \text{[#OS Needed]} \\ \vdots & & & & \end{bmatrix} \\
 \begin{bmatrix} \text{OS} & \dots & 1 & \text{[OS Position]} & \text{[#OS Needed]} \\ \vdots & & & & \end{bmatrix} \\
 \Delta t
 \end{array}
 \times
 \begin{bmatrix} \text{OS Position} \\ \vdots \\ \text{Outside Hire} \end{bmatrix}
 =
 \begin{bmatrix} \text{[#OS Needed]} \\ \vdots \\ 0 \end{bmatrix}$$



input layer hidden layer 1 hidden layer 2 hidden layer 3 output layer



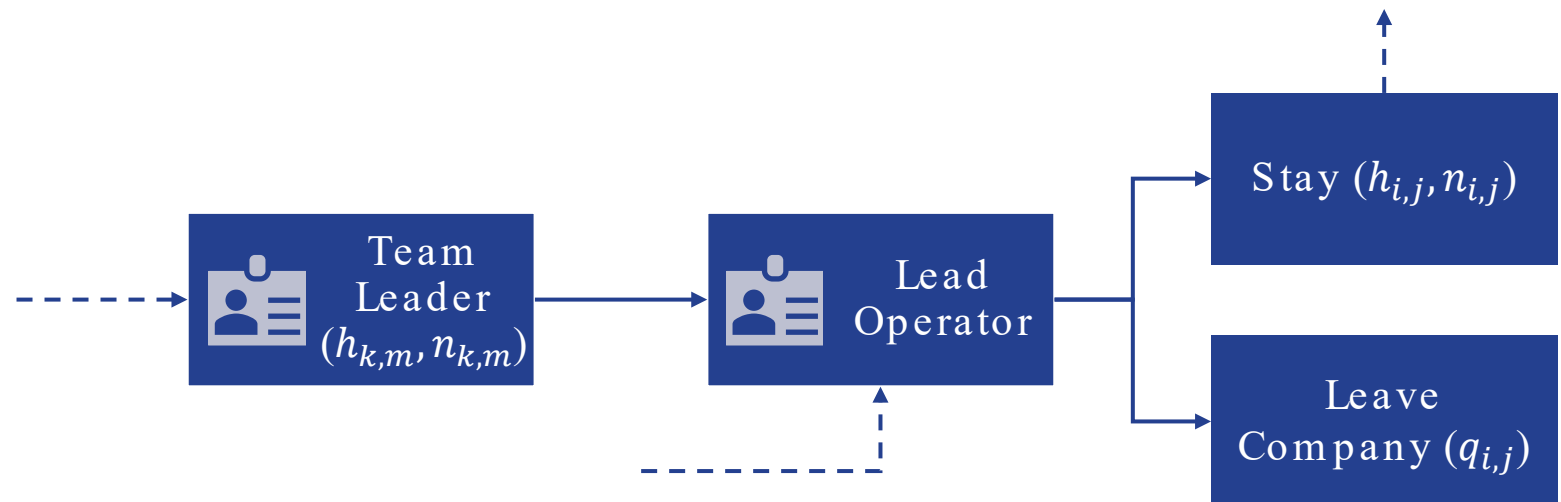
MATRIX'S ALGEBRAIC CONTENTS

$$\begin{bmatrix} OS & \cdots & \vdots \\ \vdots & \ddots & \vdots \\ & \cdots & OH \end{bmatrix} = \begin{bmatrix} h_{1,1} & \cdots & h_{1,n} \\ \vdots & \ddots & \vdots \\ h_{n,1} & \cdots & h_{n,n} \end{bmatrix}$$

$$h_{i,j} = f_{i,j} \cdot h_{k,m} \left(\frac{n_{k,m}}{n_{i,j}} \right) + q_{i,j}$$

where $\sum_j f_{i,j} = 1$

Traversed Lead Operator =
 Probability LO from TL
 • # Traversed TL
 • Weighting
 + Separations



CONTENTS FROM GRAPH

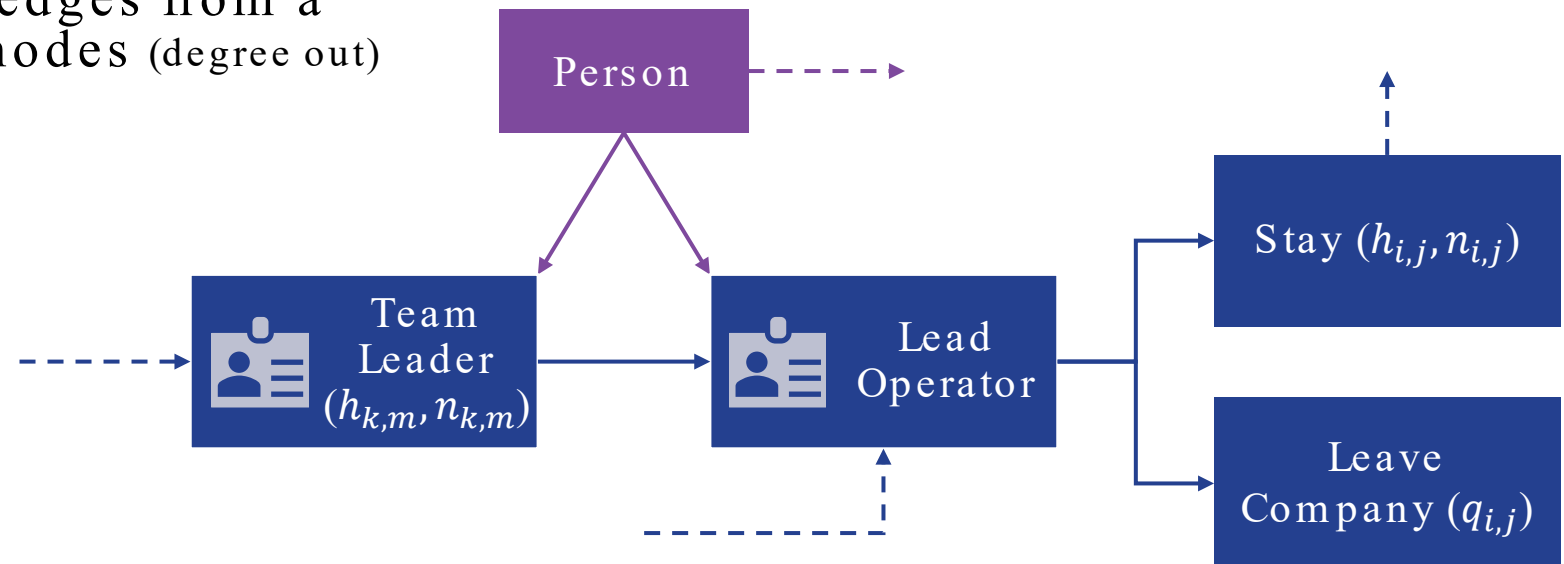
Traversed = count all edges person to a position (degree in)

Prob. = data/weight of source position

Weighting = count edge person to a position with no further positions

Separations = count edges from a position to external nodes (degree out)

Traversed Lead Operator =
Probability LO from TL
 • # *Traversed TL*
 • *Weighting*
 + *Separations*





CONTENTS FROM GRAPH

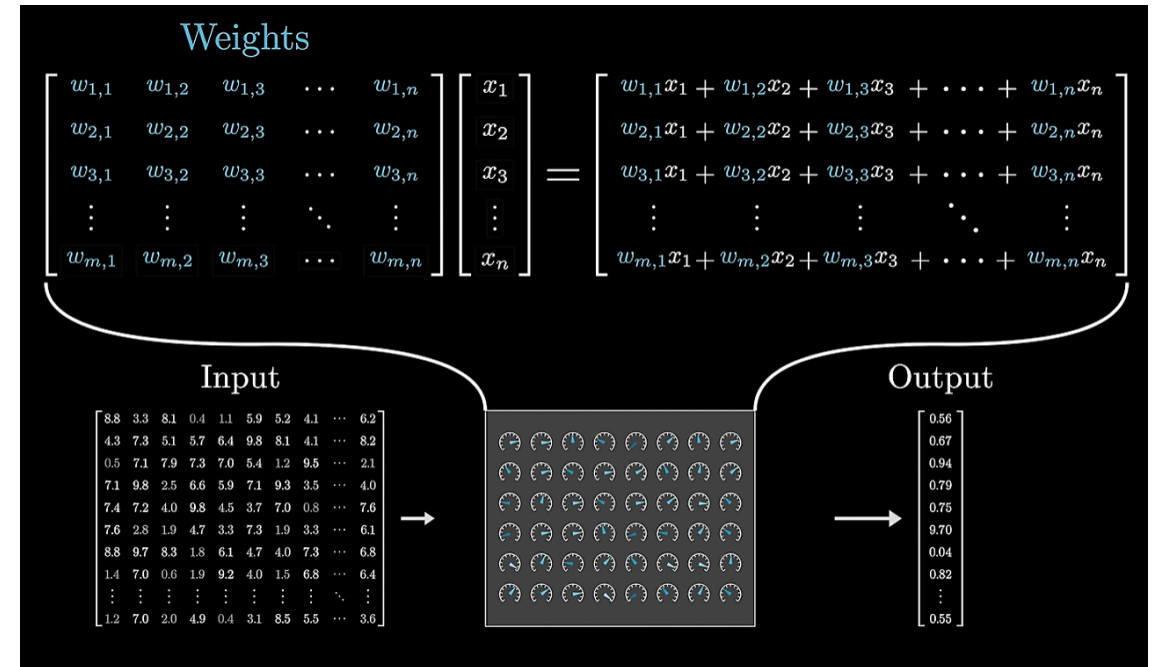
Traversed = count all edges person to a position (degree in)

Prob. = data/weight of source position

Weighting = count edge person to a position with no further positions

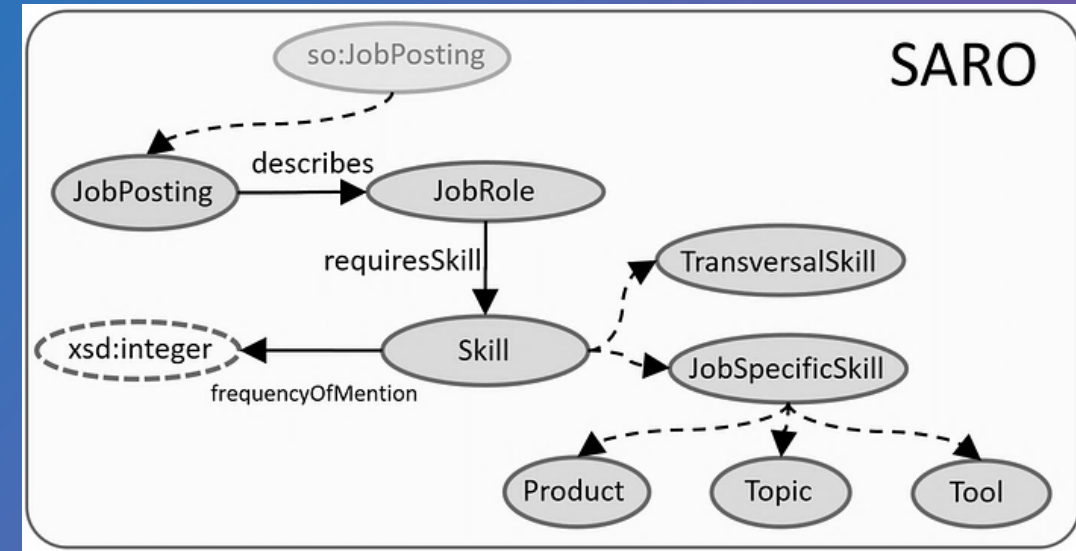
Separations = count edges from a position to external nodes (degree out)

$$\begin{bmatrix} \text{OS} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \text{OH} \end{bmatrix} \times \begin{bmatrix} \text{OS Position} \\ \vdots \\ \text{Outside Hire} \end{bmatrix} = \begin{bmatrix} \# \text{OS Needed} \\ \vdots \\ 0 \end{bmatrix}$$



ONTOLOGY

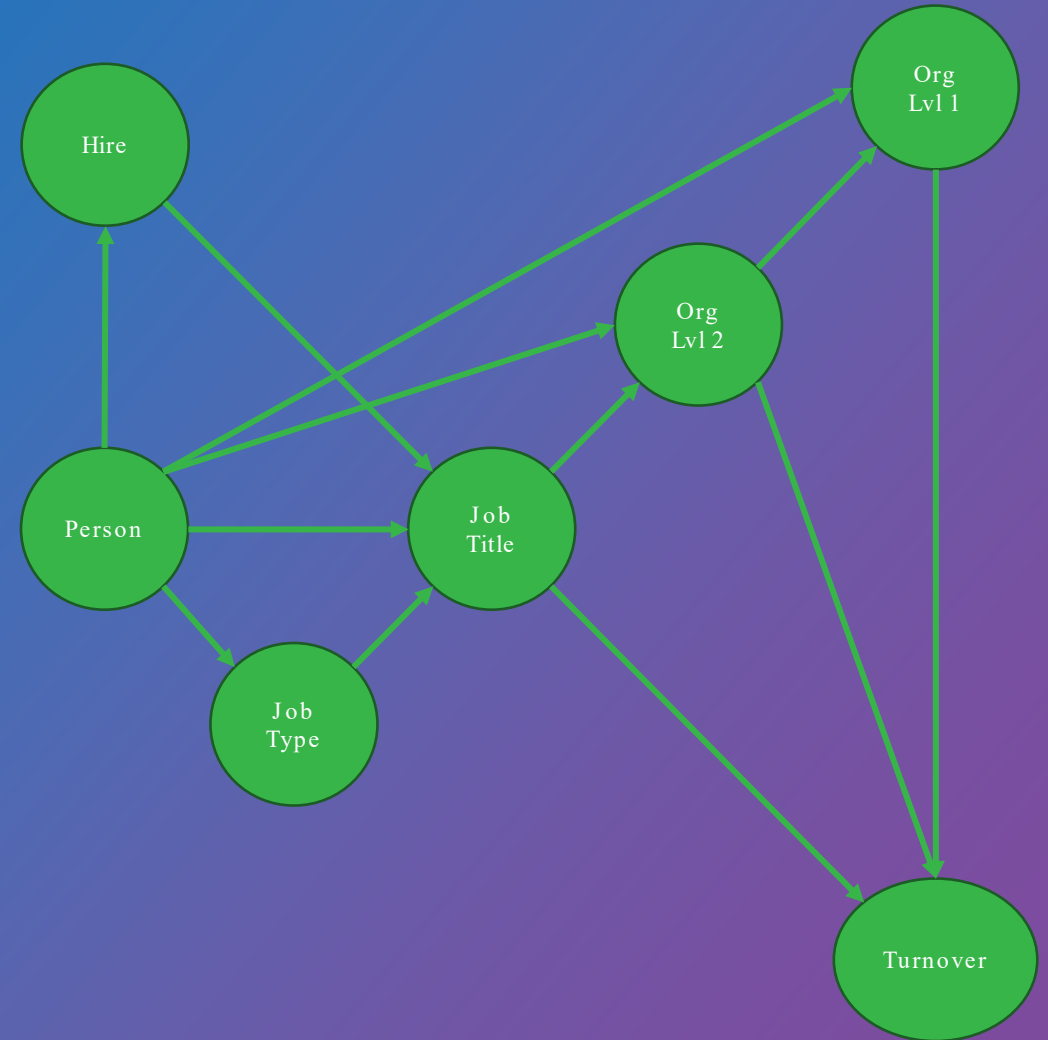
- Data quality
- Entity lineage & reproducibility
- Maintain context
- Data extraction
 - Subgraphs
 - Weights
 - ML tasks



Skills & Recruitment Ontology by the European Data Science Academy

ONTOLOGY

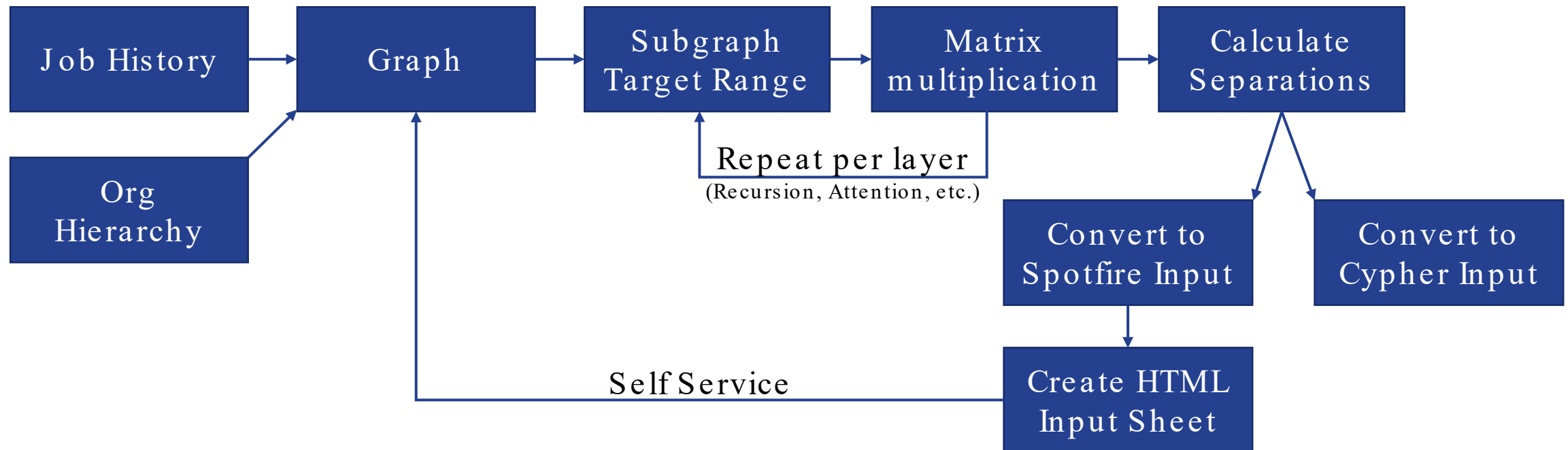
- Directed – shows movement
- Multiple indexed edges – each edge indexed to a person & time
- Density – is everything connected to everything?
- Adjacency Matrix – weighted by the attribute of investigation
 - Absolute number
 - Percentage/Probability



PROCESS FLOW

Subgraph by Edge Properties

- Edge effective date
- Status as past, current, or future position



BACKFILL PROJECTION

Output Facts

- How many people to be trained?
traversing a position
- What experience will people come with?
traversing a path (edge)

User Adjustable Weights

- Expected turnover by role
- Expected backfill source
- Required # of people by role

je	Needed Number of People	turnover
4		0.000000
1		0.000000
1		0.000000
1		0.000000
2		0.000000
1		0.000000
1		0.000000

source	target	Probability of Move
jobCode_0018	jobCode_0003	0.148148
jobCode_0025	jobCode_0005	0.027778
jobCode_0037	jobCode_0005	0.027778
jobCode_0077	jobCode_0005	0.027778
jobCode_0129	jobCode_0005	0.027778
jobCode_0255	jobCode_0005	0.027778
jobCode_0040	jobCode_0008	0.015385



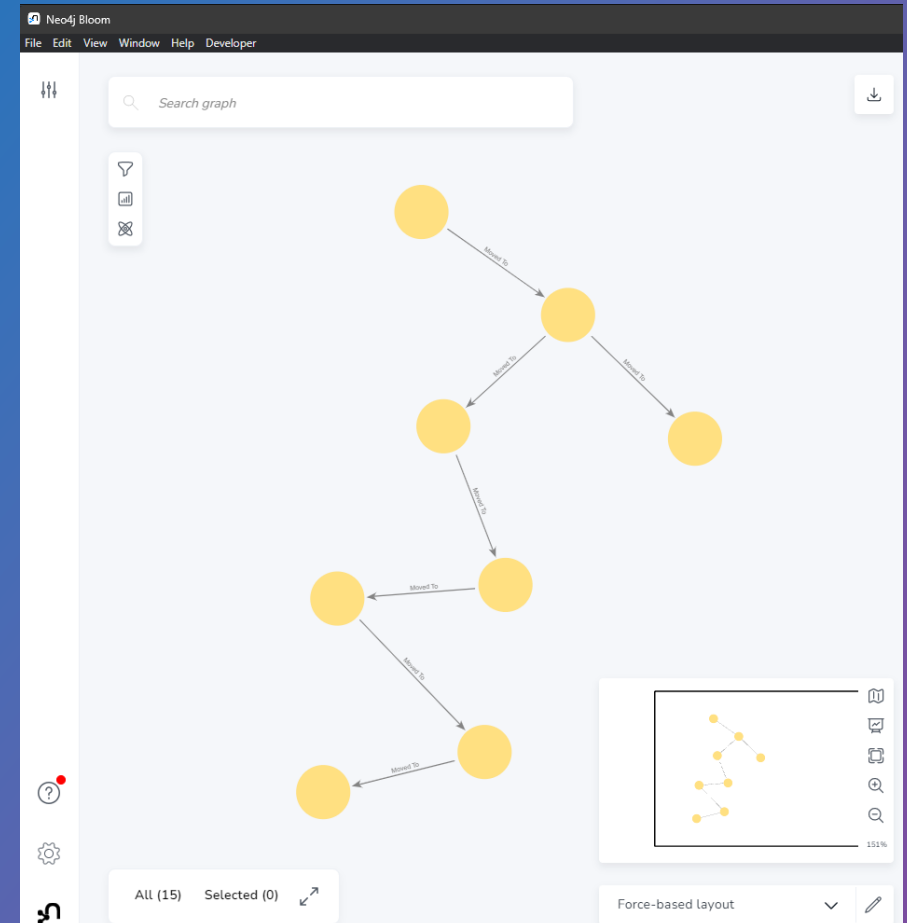
ANALYTICS INTERROGATION

Neo4j

Path from role A to role B

- What are the quickest ways to reach role B?
Shortest path algorithm
- What is the most common role on possible paths?
Betweenness centrality algorithm

```
MATCH p = allShortestPaths((n:jobCode
{value:'Role A'})-[*..8]->(m:jobCode
{value:'Role B'}))
WHERE all(r in relationships(p)
WHERE r.edge_key='...') RETURN p
```



SELF SERVICE INTERROGATION

BI TOOL (SPOTFIRE)

Purple: career advancement

- Motivation to stick with a difficult start

Blue: more likely to quit

- Known difficult starting position
- Isolate professional retention difficulties

Self Service: Noun-Verb-Noun Query/Filter

Source Noun	Verb	Target Noun
<input type="checkbox"/> employeeType		<input type="checkbox"/> employeeType
<input type="checkbox"/> HireSource		<input checked="" type="checkbox"/> jobCode
<input checked="" type="checkbox"/> jobCode		<input type="checkbox"/> orgLevel1Code
<input type="checkbox"/> orgLevel1Code	<input checked="" type="checkbox"/> Moved To	<input type="checkbox"/> orgLevel2Code
<input type="checkbox"/> orgLevel2Code	<input type="checkbox"/> Reports To	<input type="checkbox"/> orgLevel3Code
<input type="checkbox"/> orgLevel3Code	<input type="checkbox"/> Works As	<input type="checkbox"/> orgLevel4Code
<input type="checkbox"/> orgLevel4Code	<input type="checkbox"/> (Empty)	<input checked="" type="checkbox"/> TerminationReason





Graphs are
present
everywhere



Matrix
↔
Graph



AI / ML / NN
is a stack of
matrix
multiplication

- Integrate training plan
- Project hours of training required
- Create a trainer demand forecast
- Continue to optimize ontology



Conference
are sources
of inspiration



Even
inspiration is
a graph

“Gartner predicts that by 2025, graph technologies will be used in 80% of data and analytics innovations”

RECAP & UP NEXT

Thank You

The information in this presentation is the property of Hemlock Semiconductor Operations LLC (the "Company"), provided under agreement to WIT Solutions Inc. While the Company believes that information contained in this report is accurate and fully up to date, the Company does not guarantee or warranty the accuracy or completeness of information provided herein. The Company reserves the right to make improvements, corrections and/or changes to this report in the future. **THE COMPANY DOES NOT MAKE ANY WARRANTY OR REPRESENTATION, EXPRESS OR IMPLIED, WITH RESPECT TO THE UTILITY OR COMPLETENESS OF THE INFORMATION PROVIDED TO YOU AND SPECIFICALLY DISCLAIMS THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE.**