GRAPH + AI SUMMIT

organized by TigerGraph

Graph Data Science with TigerGraph Graph Algorithms

Ivan Portilla
AI Leader, IBM
portilla@gmail.com

Claire McCollough
Claire.Mccollough@colorado.edu

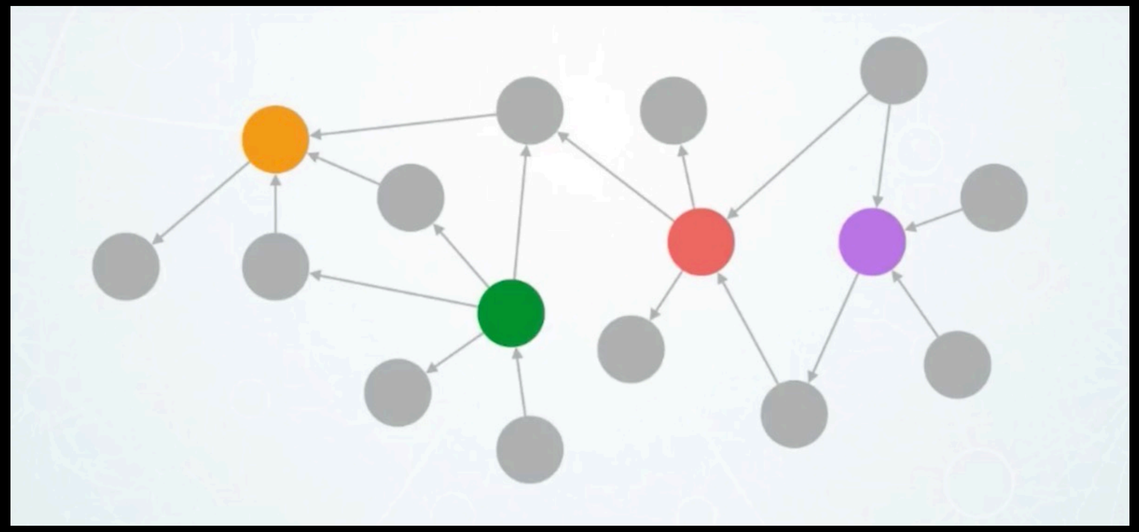
https://github.com/jiportilla/gsdl-101

Graph Data Science with TigerGraph Graph Algorithms

Ivan Portilla
AI Leader, IBM
portilla@gmail.com

Claire McCollough
Alteryx ACE
Claire.Mccollough@colorado.edu

May 2022



https://github.com/jiportilla/gsdl-101

Let me tell you a story...

https://www.youtube.com/watch?v=ifgf6bZhxiE

https://medium.com/@sweetmantech/hilton-ibm-watson-7c5f5f1a611

Today’s presentation

Graph Databases
Graph Data Science
Demonstration

Graph Data Science Featured Services

https://www.tigergraph.com

https://docs.tigergraph.com/gsql-ref/current/intro/intro

TigerGraph Graph Data Science Library

Insights at Scale with Graph Algorithms for Machine Learning.

Enterprise graph analytics and graph-native machine learning at scale.

Graph algorithms are used to compute metrics for graphs, nodes, or relationships.

Out-of-the-box Graph Data Science and Graph Machine Learning.

Watson Studio

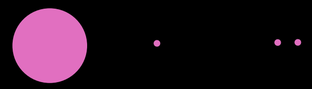
Build, run and manage AI models.

Prepare data and build models anywhere using open-source code or visual modeling.

Watson NLU

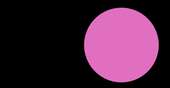
Powerful advanced text analytics for your data

https://www.ibm.com/cloud/watson-studio



Graph Databases

Graph Data Science



Demonstration

What is a graph?

A graph is ...

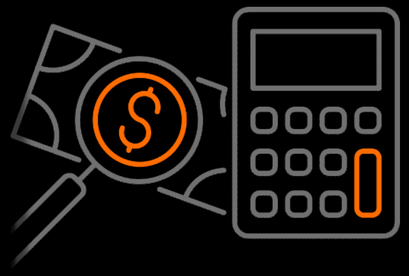
... a network of real-world entities (objects, items, or concepts) and the relationship between them.

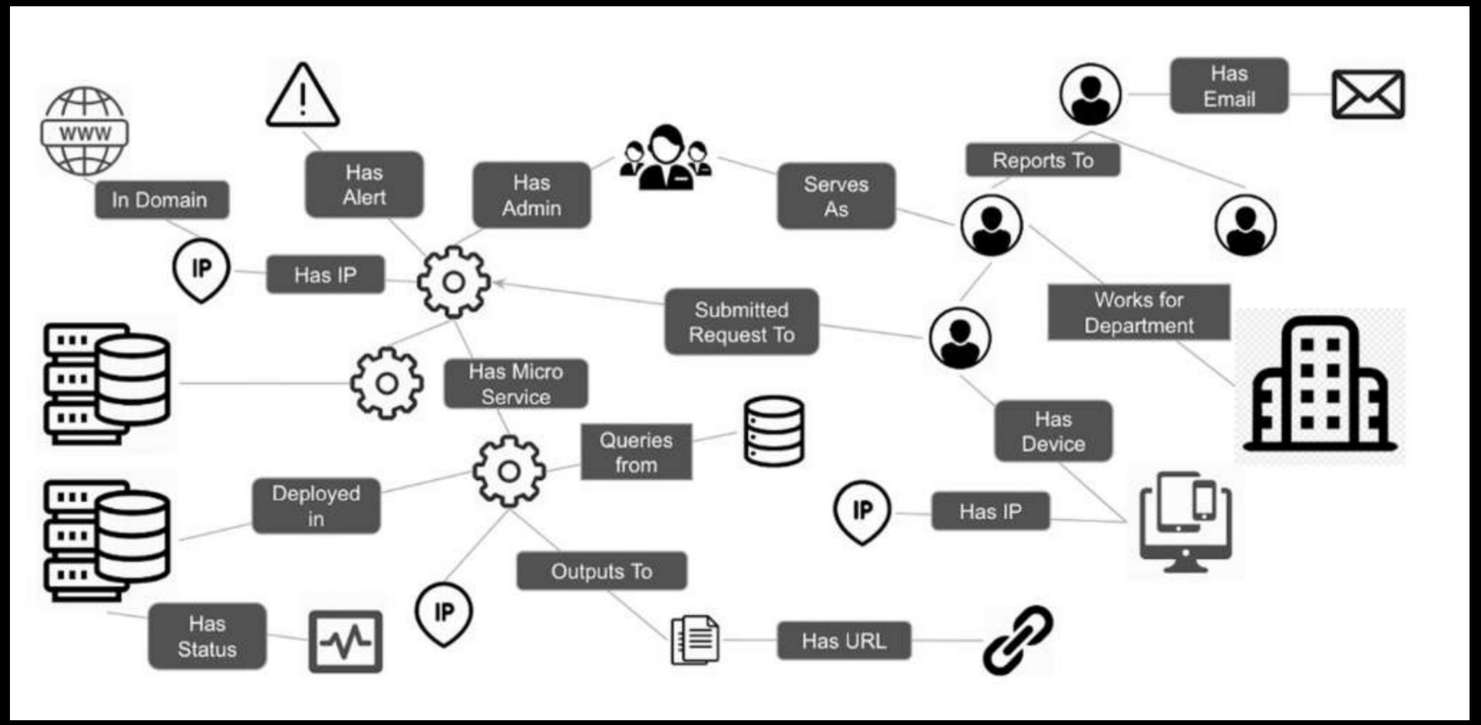
A graph database stores data with its relationships

https://www.tigergraph.com/blog/what-is-a-graph-database-and-why-should-you-care/

Why do we want graphs?

Business face the threat of Cybersecurity attacks...

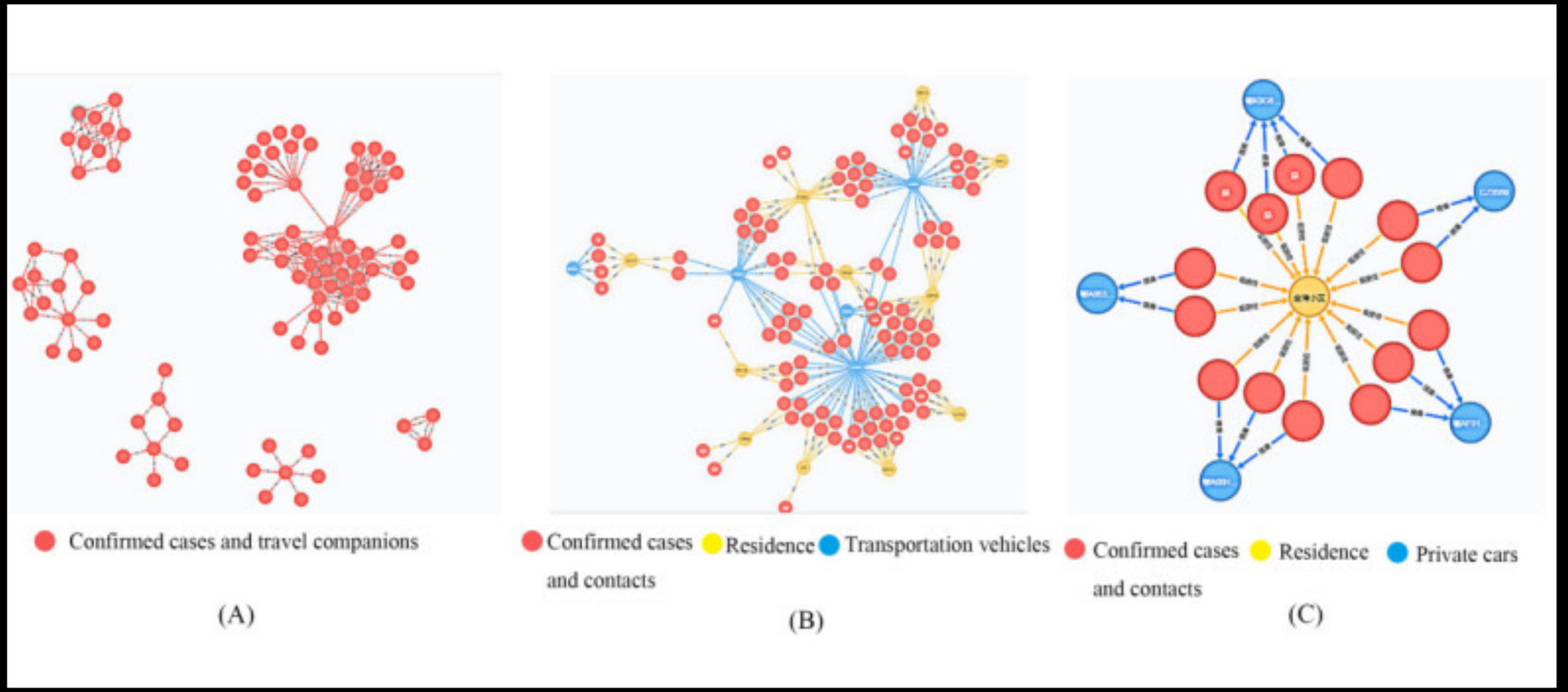




https://www.tigergraph.com/solutions/cybersecurity-threat-detection/

Why do we want graphs?

Covid-19 contact tracing...



https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7837510

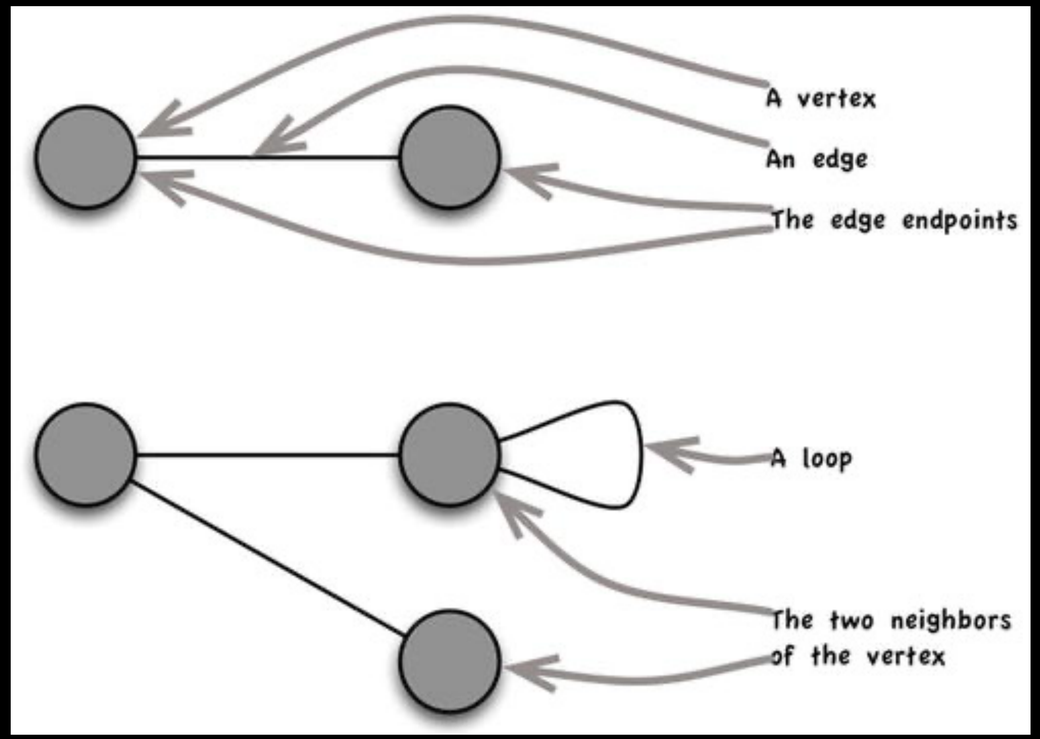
Graph Components

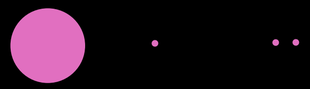
Node (Vertex)

- The main data element from which graphs are constructed

Relationship (Edge)

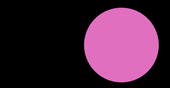
- A link between two nodes. Has:
  - Direction
  - Type
- *A node without relationships is permitted.*
- *A relationship without nodes is not*





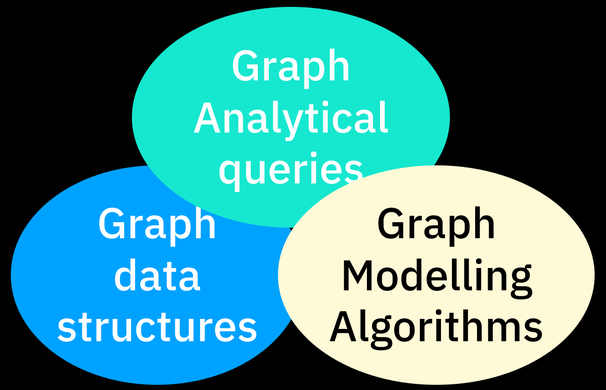
Graph Databases

Graph Data Science



Demonstration

What is *Graph* data science?



A branch of data science which employs graph data structures, graph analytical techniques and data modeling

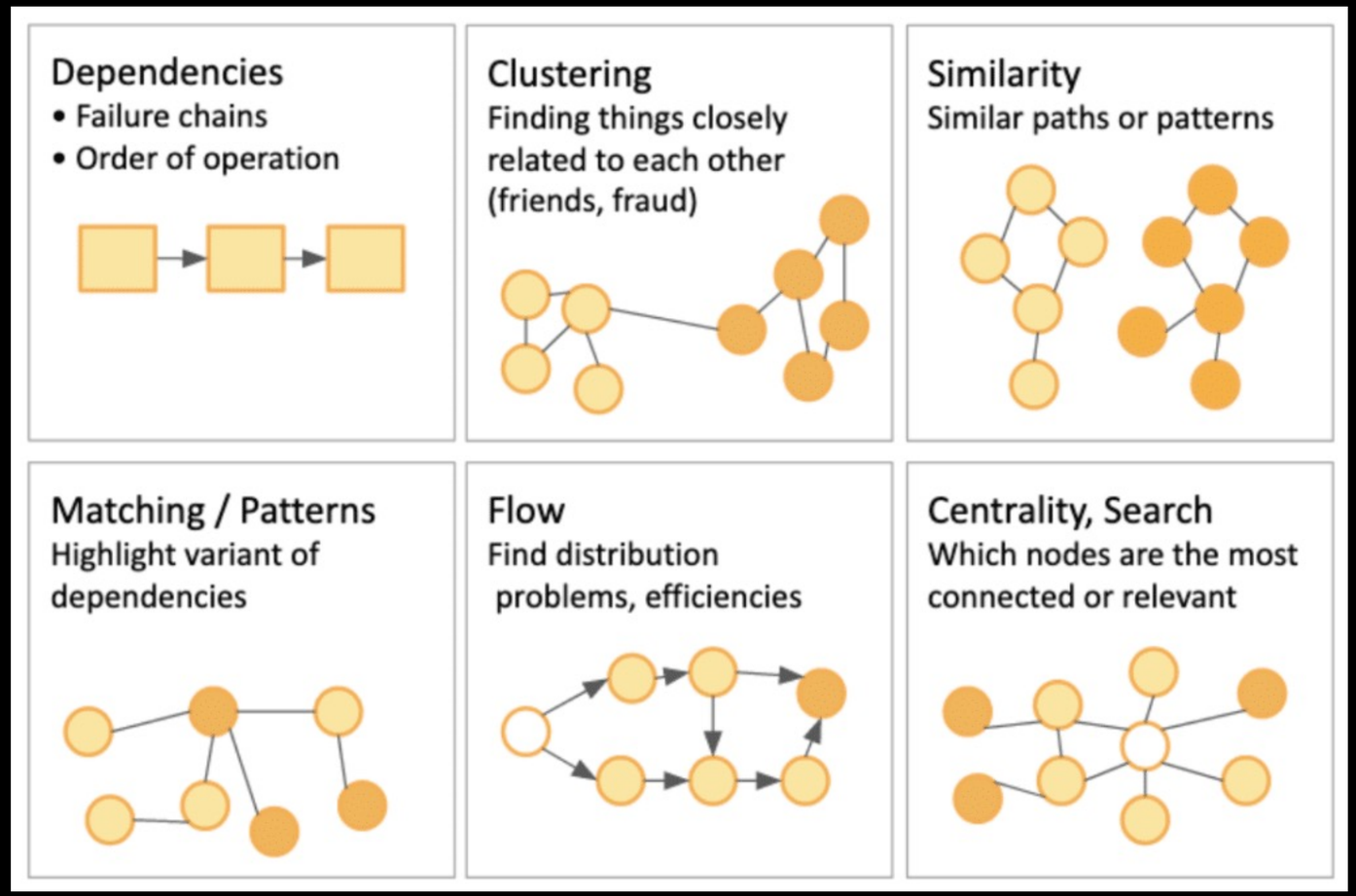
Includes several analytical methods: simple statistics-gathering queries, algorithms and machine learning methods which form predictive models.

Data scientists using relationships to answer business questions.

https://www.tigergraph.com/blog/graph-data-science-library/

TigerGraph Graph Data Science

50+ algorithms: Dependencies, Clustering, Similarity, Matching / Patterns, Flow, Centrality, and Search



https://www.tigergraph.com/graph-data-science-library/

https://github.com/tigergraph/gsql-graph-algorithms

GDSL Algorithms & Use Cases

Centrality (Search) - Determines the importance of distinct nodes in the network

Community (Cluster)- Evaluates how a group is clustered or partitioned

Classification -assigns a label to a node based on established conditions

Node Embeddings - Vector representations of properties of vertices in a graph for ML

Path - Find the optimal path or evaluates route availability and quality

Similarity - Evaluates how alike individual nodes are

Influencers in an organization → Centrality

Fraud detection, Cybersecurity → Community (Louvain)

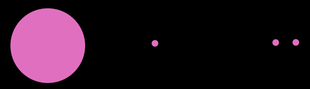
Root cause analysis → Centrality (Pagerank)

Minimum delay in telecom network → Path finding

Most Influential Health Care Provider → Community + Centrality

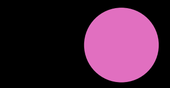
Important routers in a telecom network → Centrality

Customer classification based on purchases → Node Similarity



Graph Databases

Graph Data Science

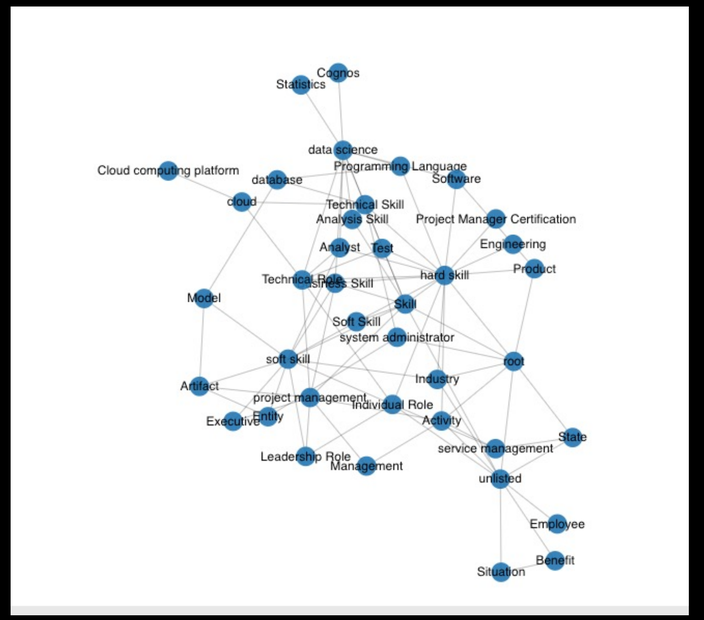


Demonstration

Demo

Demonstration

Step 1. Create the ontology data
Step 2. Run Jupyter Notebook, create TigerGraph database
Step 3. Visualize the results



Demonstration

Knowledge Graphs: Ontology
Centrality: Pagerank
Community: Louvain

Get started

1. Access TigerGraph
(or cloud sandbox, docker container, Enterprise edition)
https://www.tigergraph.com/get-tigergraph/

2. Install GDSL library
https://github.com/tigergraph/gsql-graph-algorithms

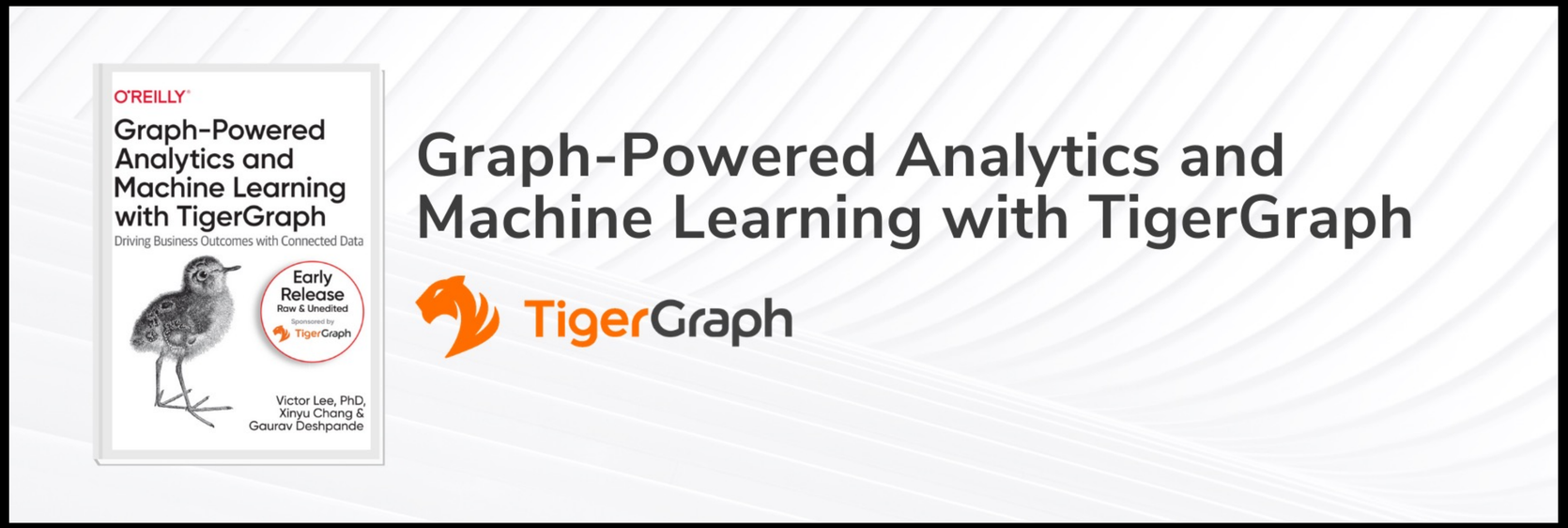
3. Watch the tutorials
https://info.tigergraph.com/graph-ai-summit-fall-session-graph-data-science-library

Q&A

Ivan Portilla

ivanp@us.ibm.com

@iportilla



https://info.tigergraph.com/oreilly-book