AML2021

ANTI-MONEY LAUNDERING & FINANCIAL CRIMES VIRTUAL CONFERENCE

Improving AML Detection and Investigation using Machine Learning and Graph Databases

Today’s Speakers

David Ronald
Product Marketing Director
TigerGraph

Steven Fuller
Senior Solutions Engineer
TigerGraph

UNODC
United Nations Office on Drugs and Crime

The United Nations Vienna 1988 Convention Article 3.1 defines ***money laundering*** as "*the conversion or transfer of property, knowing that such property is derived from any offense(s), for the purpose of concealing or disguising the illicit origin of the property or of assisting any person who is involved in such offense(s) to evade the legal consequences of his actions*".

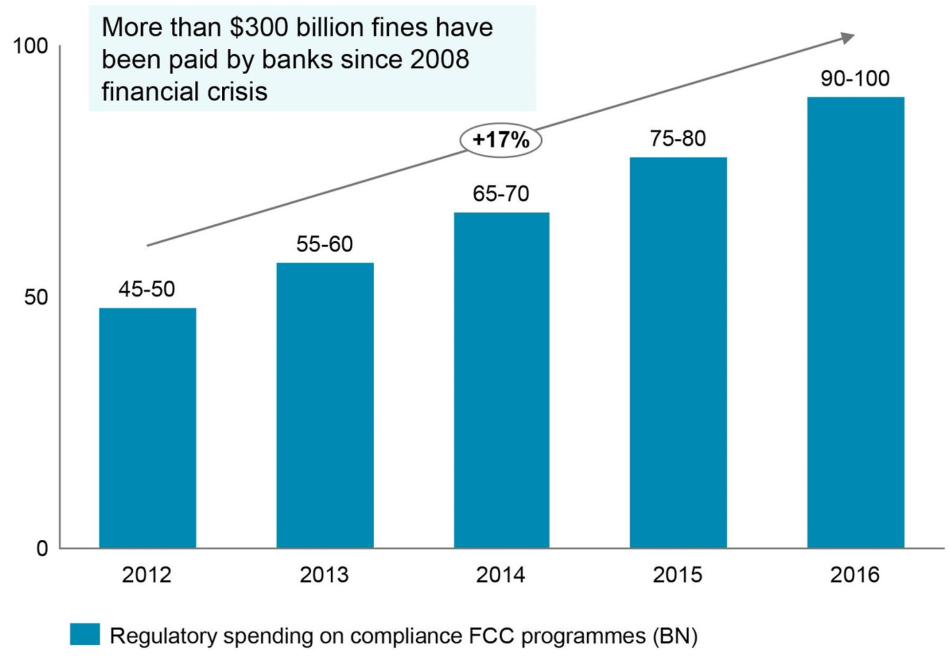
Three stages of Money Laundering:

- **Placement**: Moving funds from direct association with crimes
- **Layering**: Disguising the trail to foil pursuit
- **Integration**: Making money available to criminals from what seem to be legitimate sources

https://www.unodc.org/unodc/en/money-laundering/overview.html

Total aggregated US bank fines in 2020	**$14.21B**
Most common violation	AML breaches

https://finbold.com/bank-fines-2020/



AML Considerations

Not Small	UNODC: Money laundering is now estimated to be "2–5% of global GDP, or $800B to $2T in current USD" (>AAPL market cap)
Not Just Banks	Entirely new counterparties now: Retail, casinos, real estate, human and drug trafficking, terrorists, cryptocurrencies, and other **networks**
Not Just Risk Scoring	Investigations need deep and wide and fast network forensics: **financial crime networks are evolving into advanced persistent threat systems**
Not Just KYC	Effective investigations also now need to avoid false positives: **KNOW YOUR CUSTOMER'S CUSTOMER: KYC(C)**
Not Just Humans	Graph analytics, together with ML/AI and data science - **with humans in the loop** - together help investigators do even better

Technology Use Cases

Business Cases

- Anti Money Laundering
- Internal Fraud - Entitlements
- Credit Card & Transaction Fraud
- Identity Theft & Falsification
- Cyber - Malware
- IoT & Asset Fraud
- Audit & Compliance
- Claim, Dispute Charges
- Law Enforcement - Prosecution

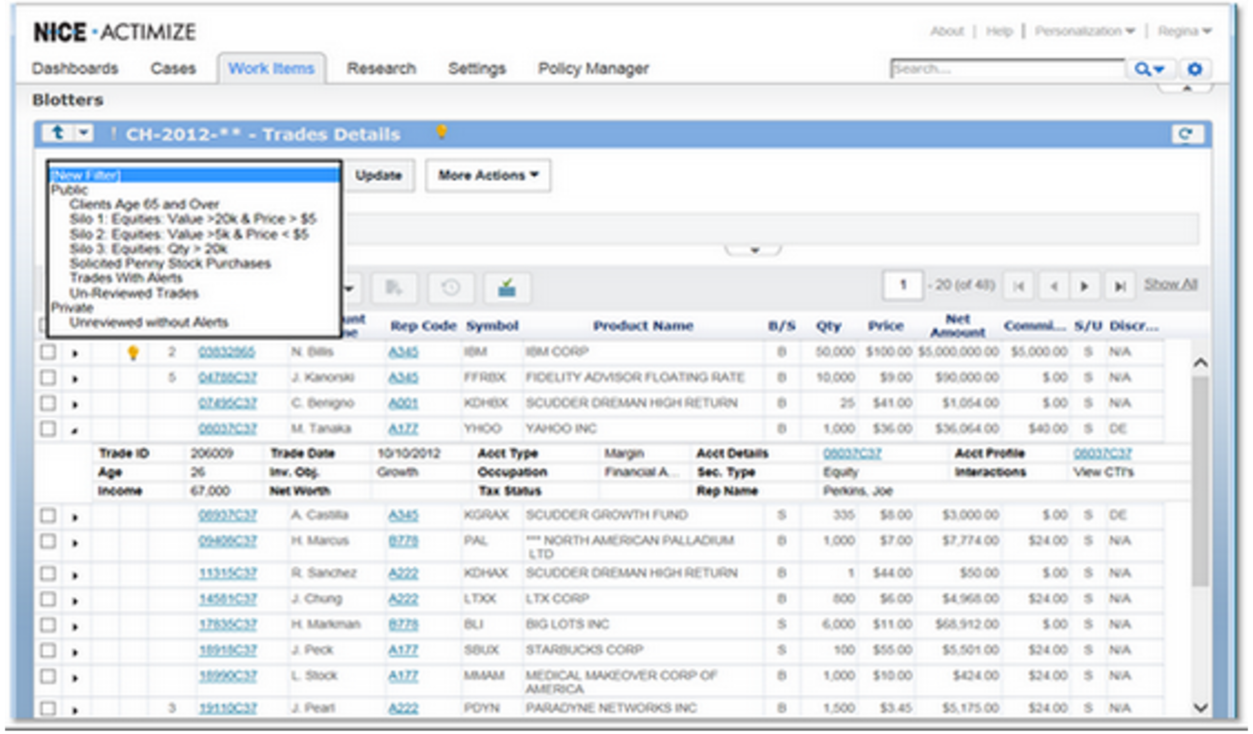
Investigation - Visualization

- Advanced Visualization
- Dependency | Network Pathing | Routing | Complex Visualization
- Clustering & Community Detection
- Geospatial 'Network Mapping'
- Real Time Data - IoT Systems
- Team-based Workbench & Investigation

ML and Analytics

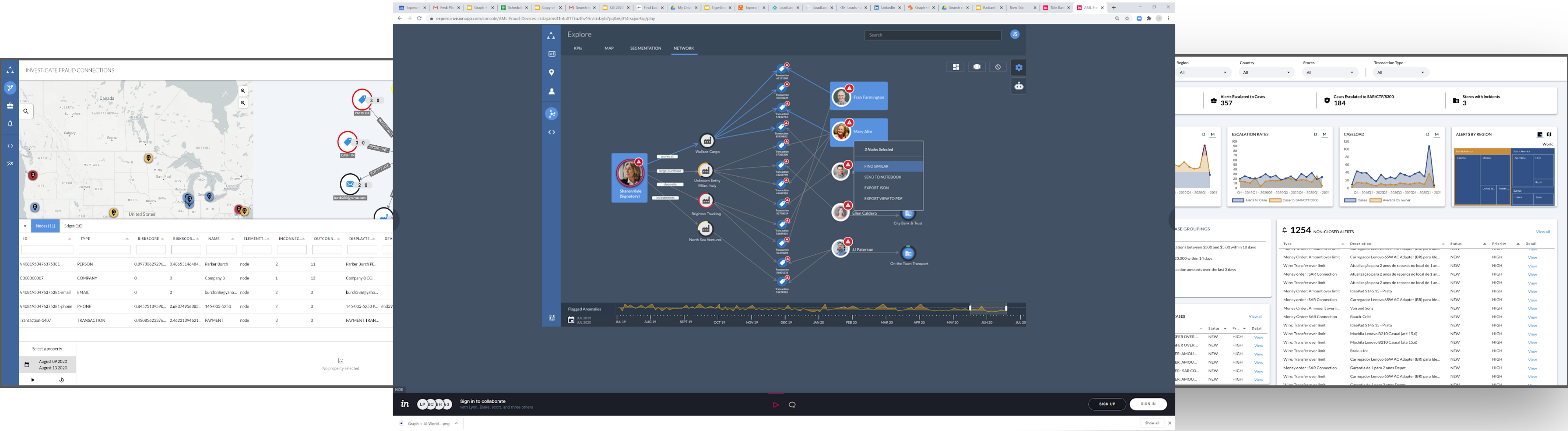
- Patterns - Recommendations
- What If - Planning & Visibility
- Predictive & Analytics
- Scoring and Risk
- Audit & Compliance - Historical
- Targeting Similarities
- Decision Tree Analytics

Limited Legacy Systems



- No depth of context:
  Mules? Sometimes.
  Fraud rings? No.
- Heavy overhead in terms of integration, operations, human-in-the-loop investigations.
- Lack of speed, scale, insight into connected data.

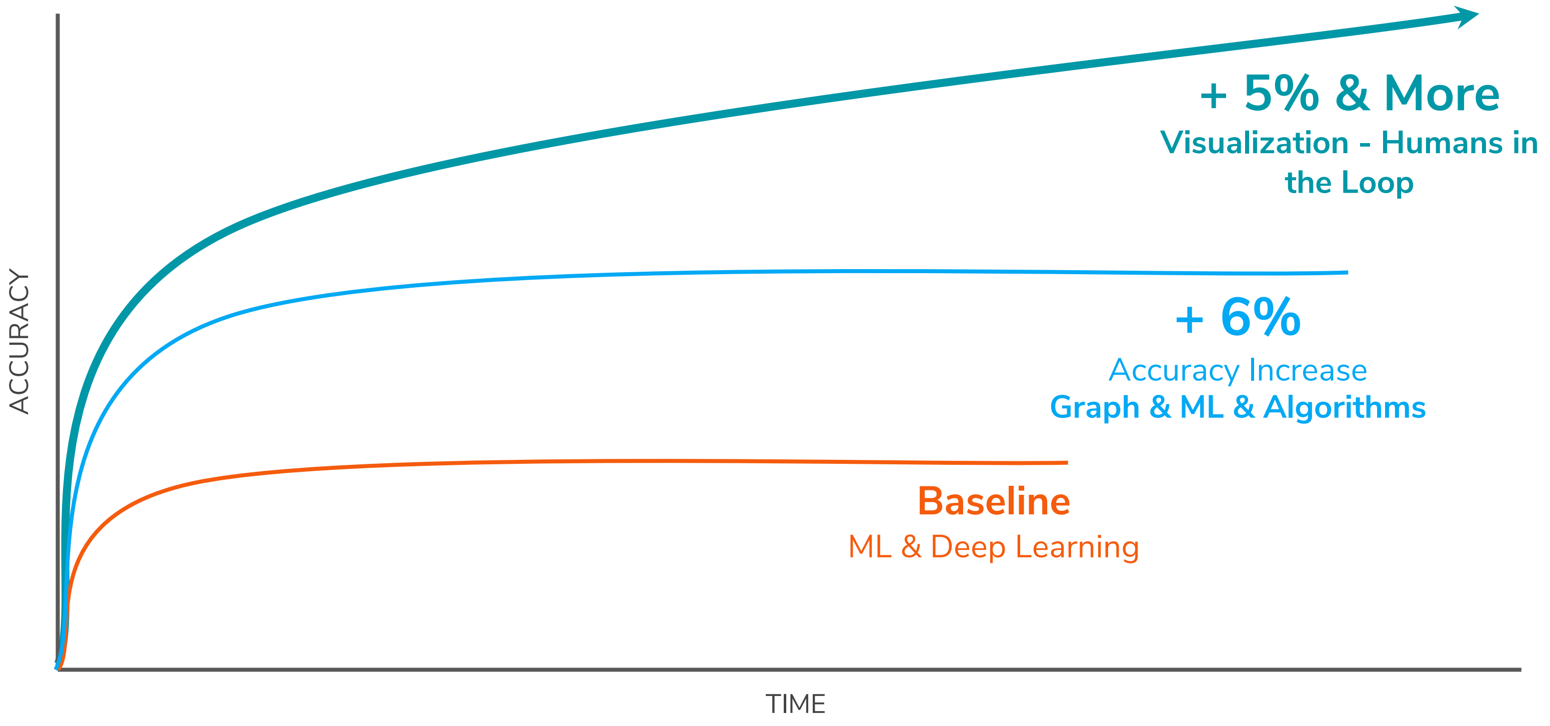
Better Investigative Insights



- Actionable investigative insights obtained using machine learning and graph databases

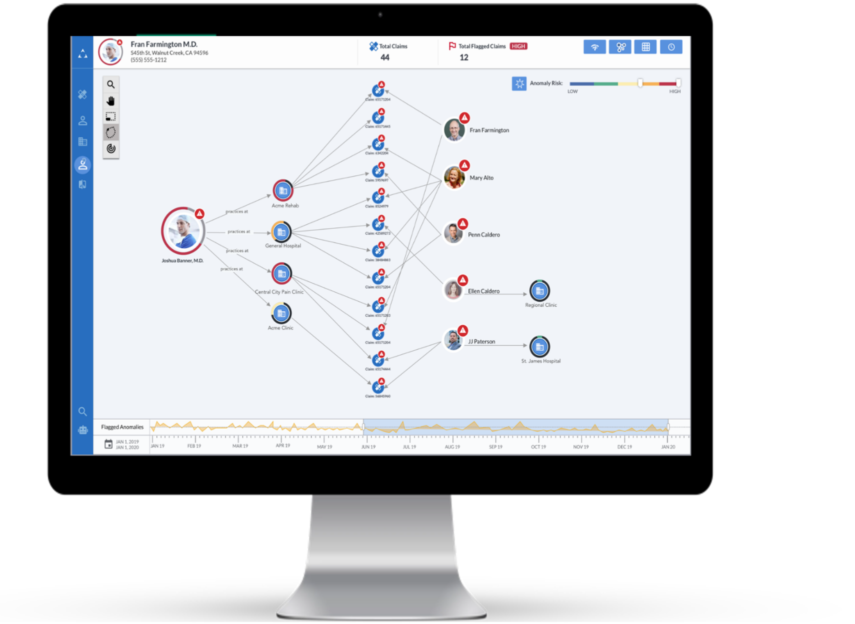
Better Analytics Accuracy

Graph + machine learning + humans = improved accuracy of AML investigations



Benefits of AML Workbench

- **Increase accuracy of alerts** - human and algorithms (eg, false positives)
- **Decrease false positives -** capture and incorporate human instincts and expertise (especially for expert users)
- **Increase speed & efficiency** - more accurate and trusted outcomes (both automated & human)
- **Improve operational efficiency** - reduce the level of effort needed to investigate alerts and create cases
- **Adapt to changes** - add new alerts and rules to address new patterns
- **Continuous improvement -** system can learn, adapt & get better over time



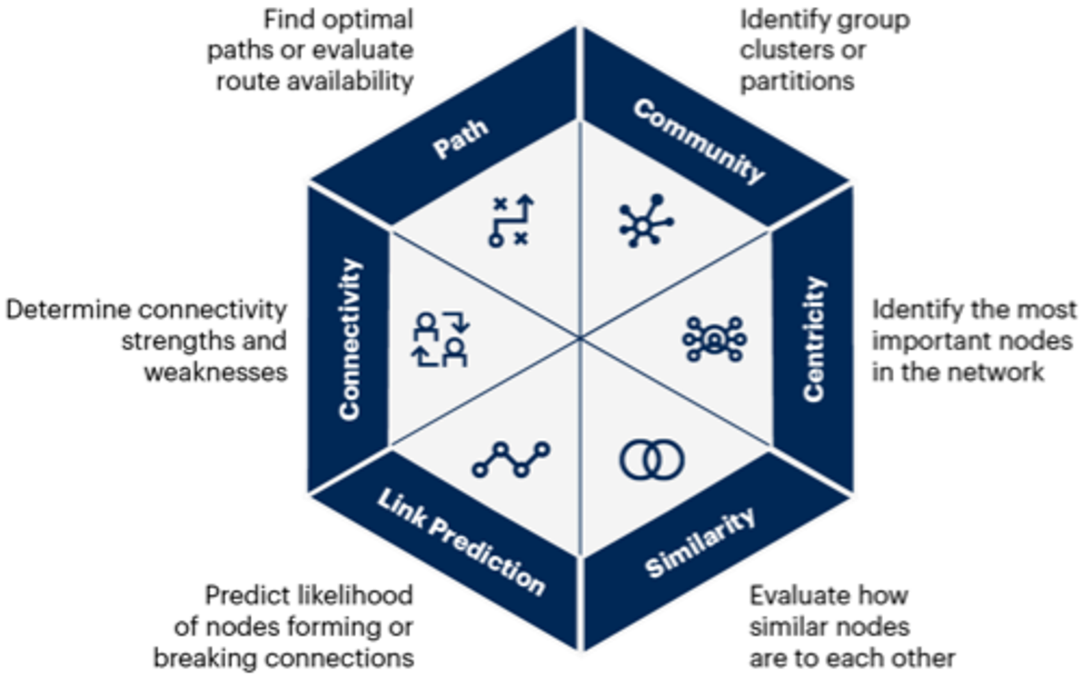
Introduction to Graph Databases

"Graph analysis is possibly the single most effective competitive differentiator for organizations pursuing data-driven operations and decisions after the design of data capture."

Gartner

Six Types of Graph Analytics

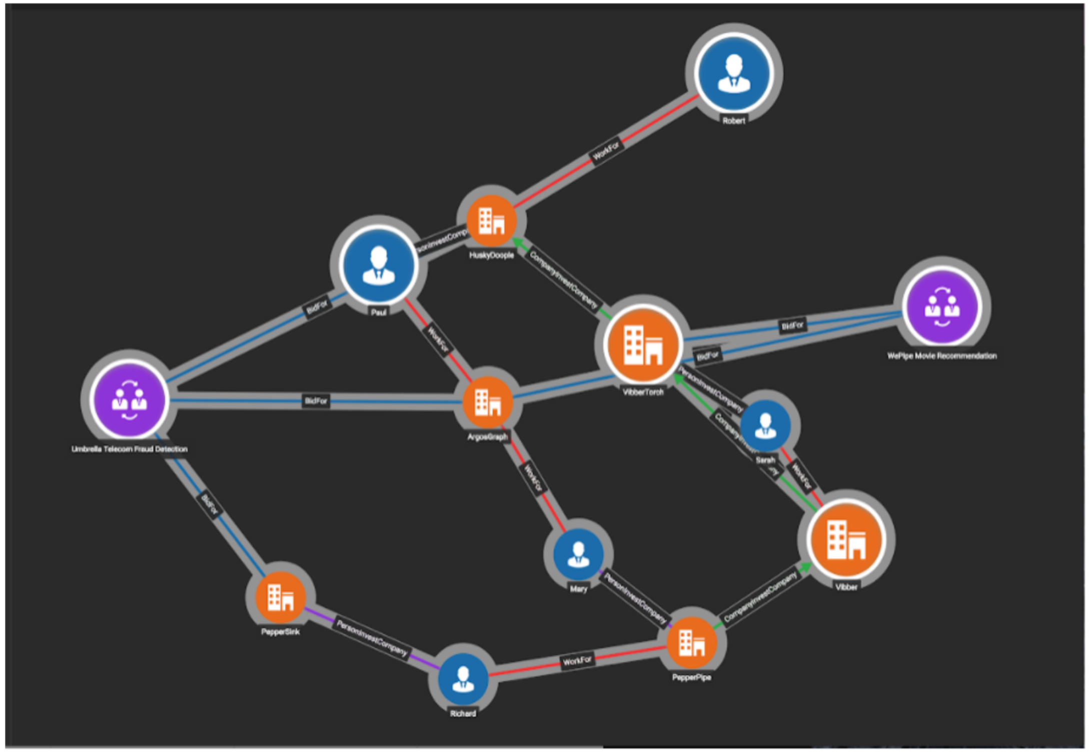
Graph can be used to analyze all sorts of relationships across all kinds of systems even beyond process or beyond the confines of individual operational models.



Source: "Understanding When Graph Technologies Are Best for Your Business Use Case", Jim Hare et al, 2020

Gartner

A Sample Graph

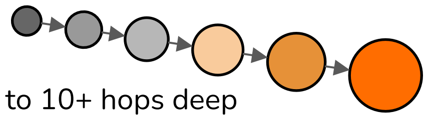
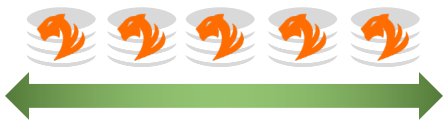
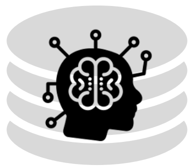


Graph databases consist of vertices and edges

- Vertices - data entities
    - for example - person, account, transaction
- Edges - the relationships between those entities
    - for example - person opens account, money moves from one account to another account

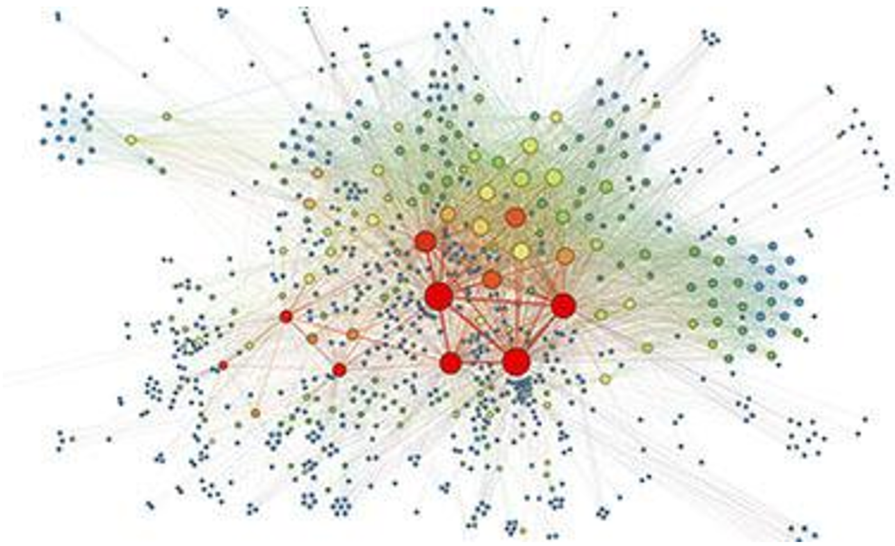
A graph stores the relationships between data entities - or can be used to uncover relationships between data entities

Why Graph Databases?

Feature	Design Difference	Benefit
Deep-Link Pattern Discovery  5 to 10+ hops deep	• Native graph, for speed and efficiency	• Uncovers hard-to-find patterns • Operational, real-time analytics
Handling Massive Scale 	• Distributed database architecture • Massively parallel processing • Compressed storage reduces footprint and messaging	• Integrates all your data • Automatic partitioning • Complete data → better detection
In-Database Analytics 	• GSQL: High-level yet Turing-complete language • User-extensible graph algorithm library, runs in-database • ACID (OLTP) & accumulators (OLAP)	• Avoids transferring data • Richer graph context • In-place machine learning

Gartner

Better AML with Graph



Richer, smarter connected data
- Uncover connections, break down silos

Deeper, more far-reaching questions
- Search far and wide faster and more easily

Enhance investigations with non-obvious information
- Connect SARs from prior investigations

Discover hidden behaviors and compare patterns
- Decide which investigations have a high probability of a successful outcome

Explainable results
- AI/ML-based results are credible and understood by humans
- Visual exploration

About TigerGraph

Connect Datasets & Pipelines

Friction-free scale out from GB to TB to PB with lowest cost of ownership

UNITEDHEALTH GROUP®

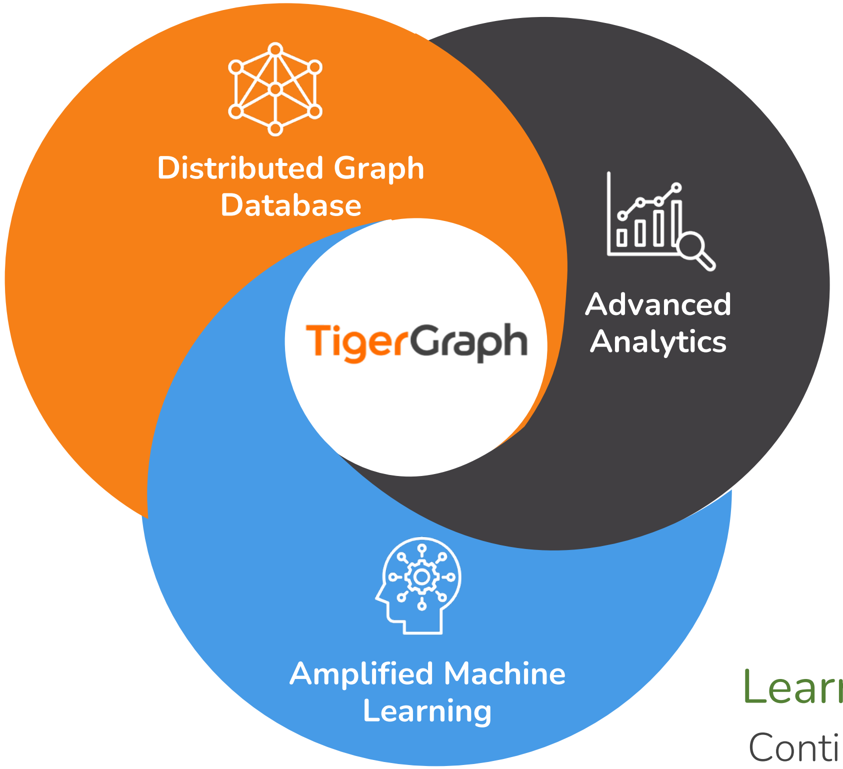
Customer 360 connecting 200+ datasets and pipelines

Fortune 50 Retailer

Item 360 for ecommerce across 100+ datasets

xandr AT&T

Identity graph connecting 50+ data pipelines



Analyze Connected Data

10-100x faster than current solutions

JAGUAR LAND ROVER

Supply chain planning: 3 weeks to 45 minutes

China Mobile International

Fraud Detection: batch to real-time for 300M calls/day

Learn from Connected Data

Continuous graph-based feature generation & training

intuit.

AI-based Customer 360 for entity resolution, recommendation engine, fraud detection

Mapping a Road to Success

Data sources:
- Product Catalog
- Transactions
- Operations
- Customers
- Social Networks
- Sales CRM

1 AML Graph

2

Link & Merge Entities
- Rule-based
- ML-based

3

Build Initial Visualization System

Human Exploratory Learning
- Get to know your business/data better
- Define useful measures and displays, particularly for fraud investigation and confirmation

Unsupervised Learning
- Run graph algorithms to characterize your data: connections, groupings, frequent patterns, outliers
- What is typical/frequent vs. atypical?

4

Use Graph+ML to Detect & Investigate Money Laundering

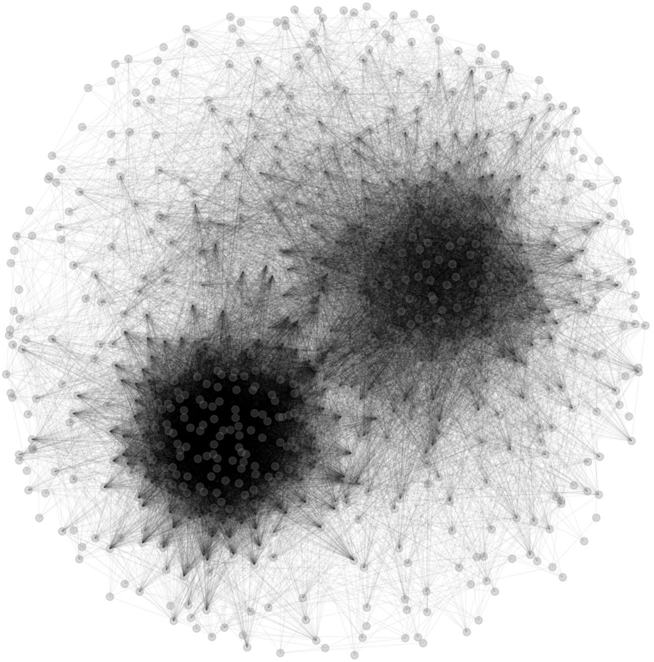
Extract Graph Features
- Measure closeness and paths to entities of interest
- Detect communities and clusters (Louvain)
- Measure similarity between entities (Cosine, Jaccard)
- Rank the influence of entities (PageRank)
- Graph Embedding (Node2Vec)

Gather Training Data

Train Your ML Model with Graph Data
- Export graph data to external ML platform
- In-database ML

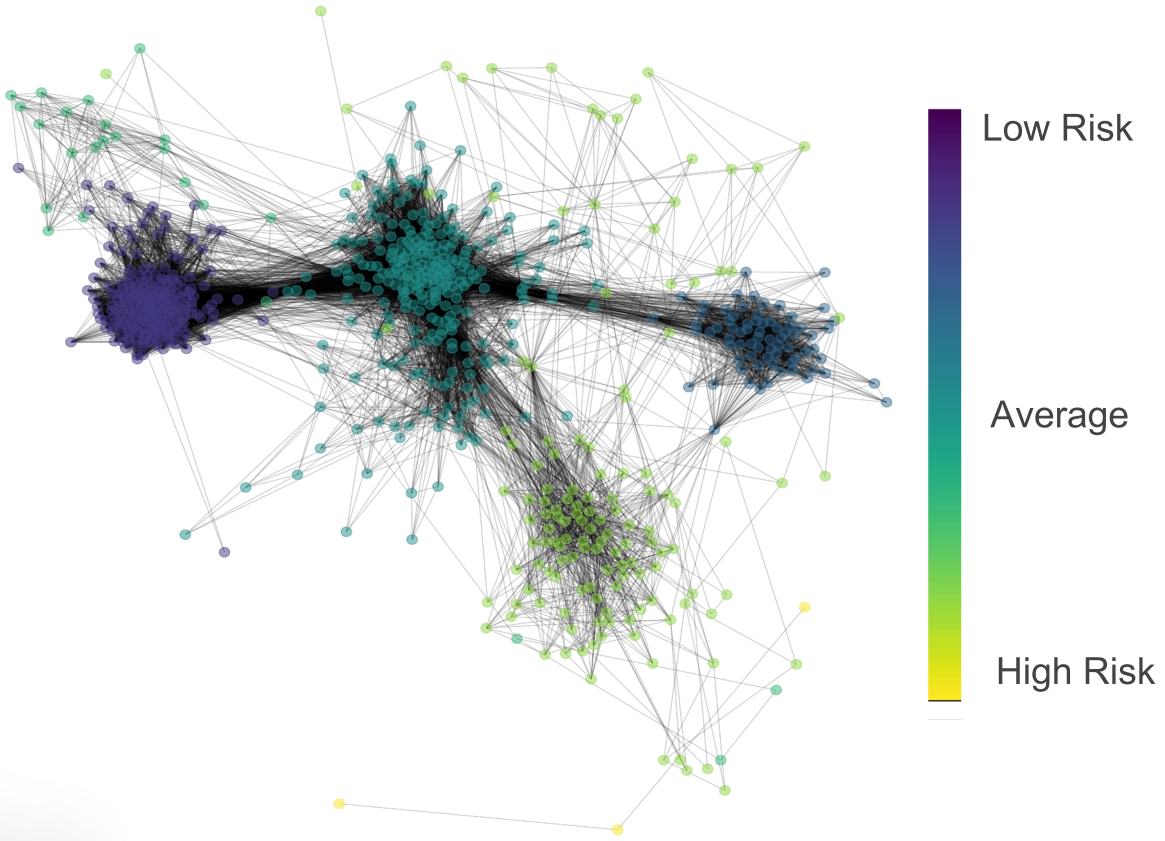
Viewing Risk in a Graph

An unstructured graph of
unrelated bank transactions



vs.

Same graph, automatically clustered by their history
Louvain Communities with graph + ML algorithms



Demonstration

Thank You

David Ronald
Product Marketing Director
TigerGraph
david.ronald@tigergraph.com

Steven Fuller
Senior Solutions Engineer
TigerGraph
steven.fuller@tigergraph.com

Thank You