EECS 6895 Advanced Big Data and AI

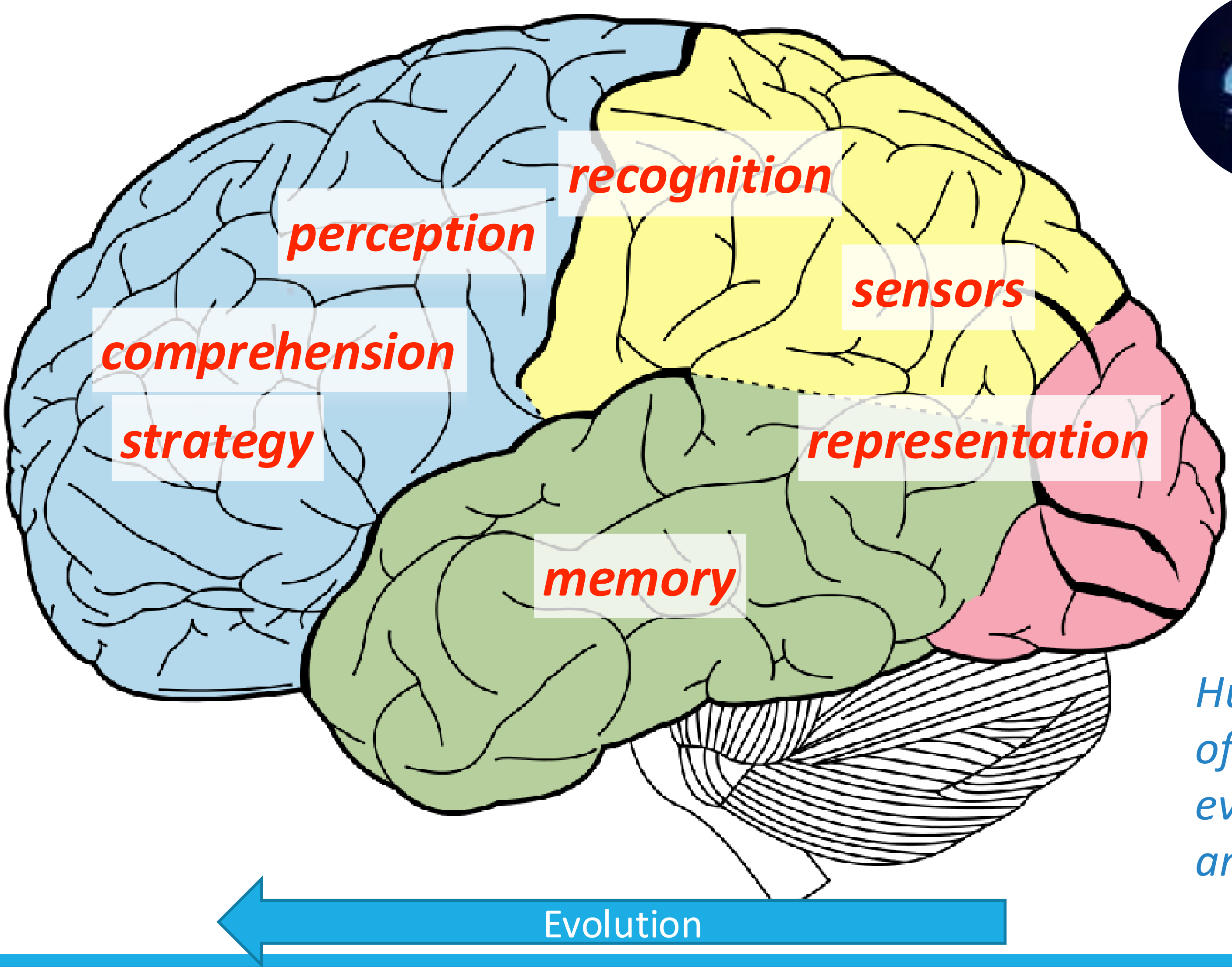
Lecture 14: Full-Brain AI

Prof. Ching-Yung Lin

Columbia University

April 29th, 2025

Ardi Platform



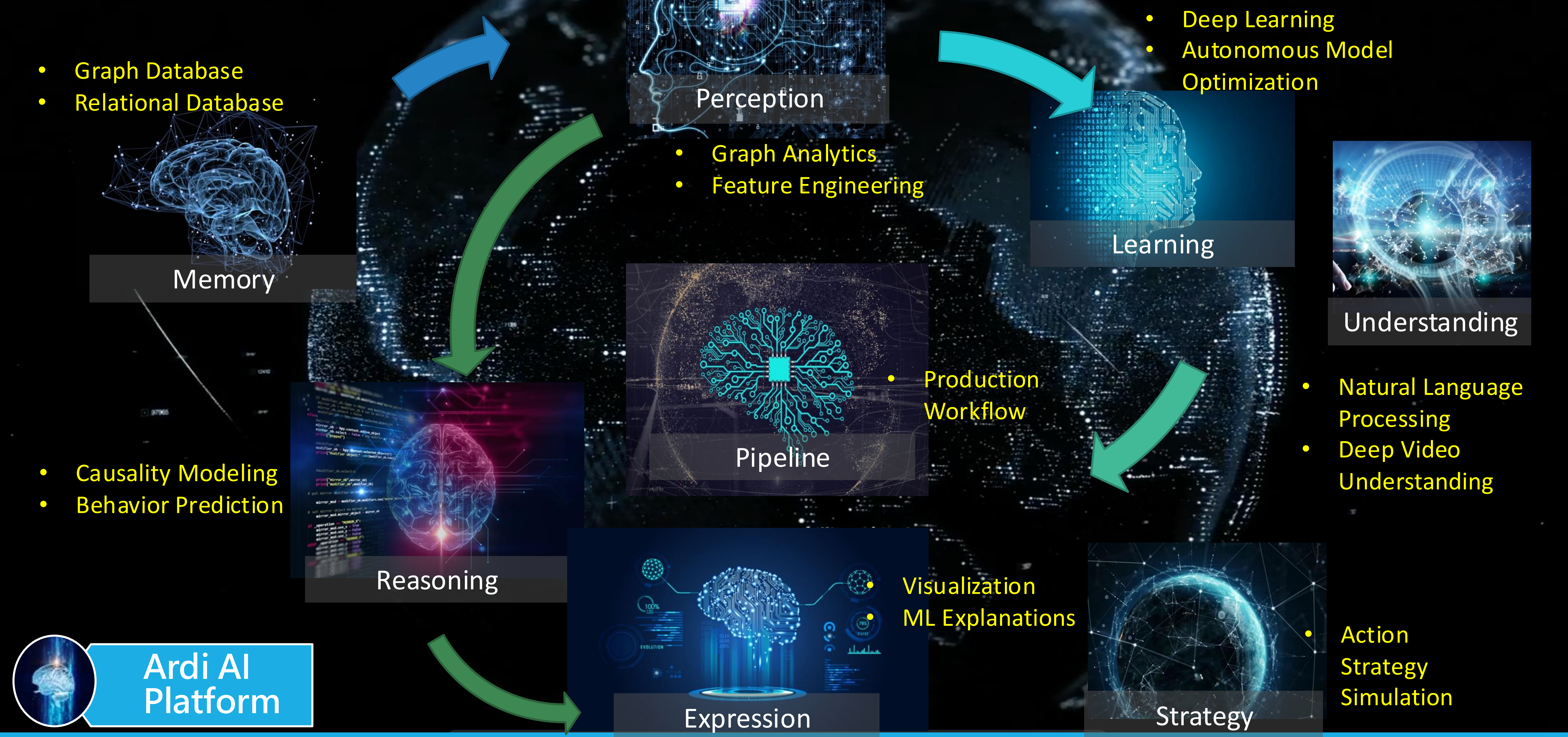
Ardi AI Platform

Contextual Analysis | Autonomous Learning

Advanced Enterprise Full-Brain AI Platform to build solutions – Scalability, Stability, and Advanced AI Technologies

Human Brain – a graph network of 100B nodes and 700T edges evolved and became smarter and smarter.

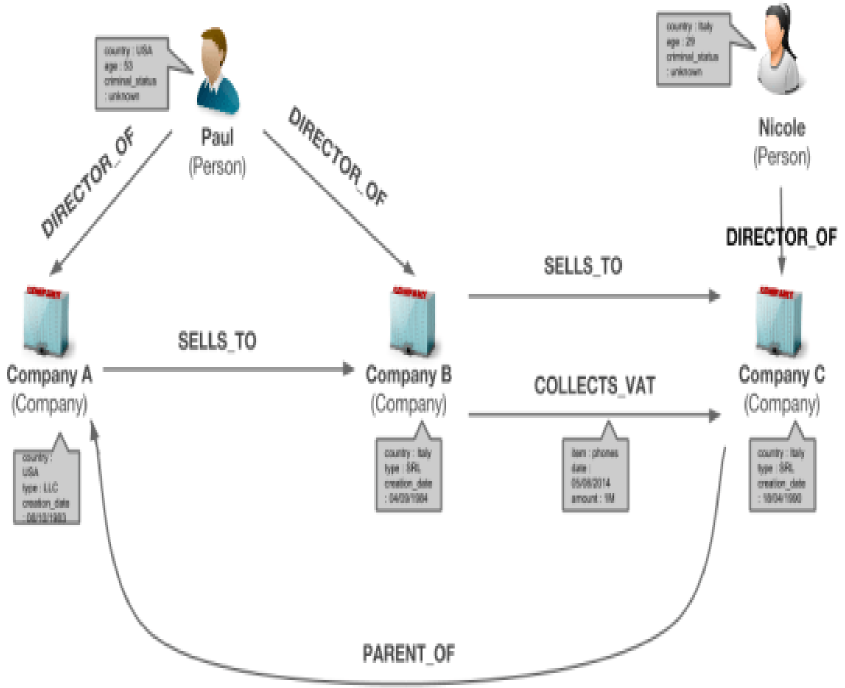
Ardi's Enterprise- Ready Functions



Three Major Reasons Why Graph makes Machine Smarter

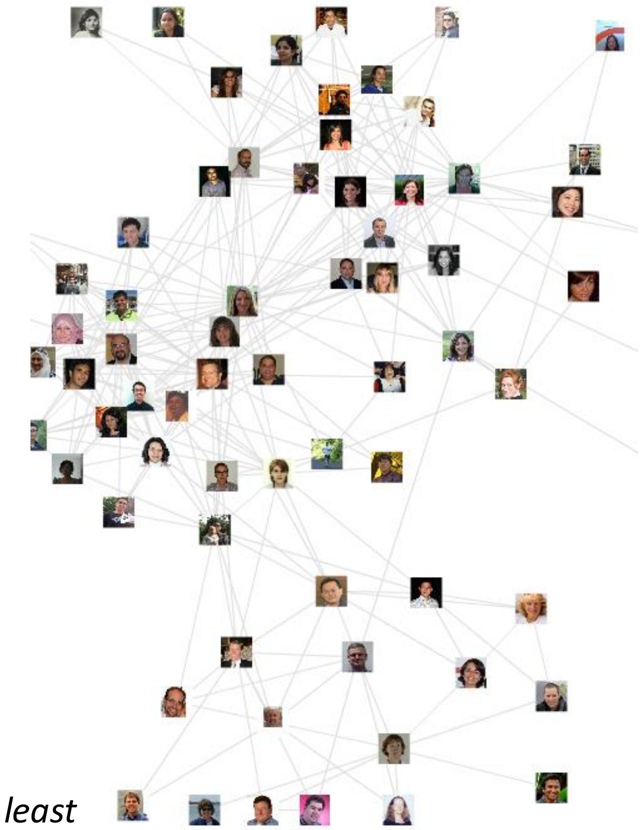
Context

Attributes and Relationships



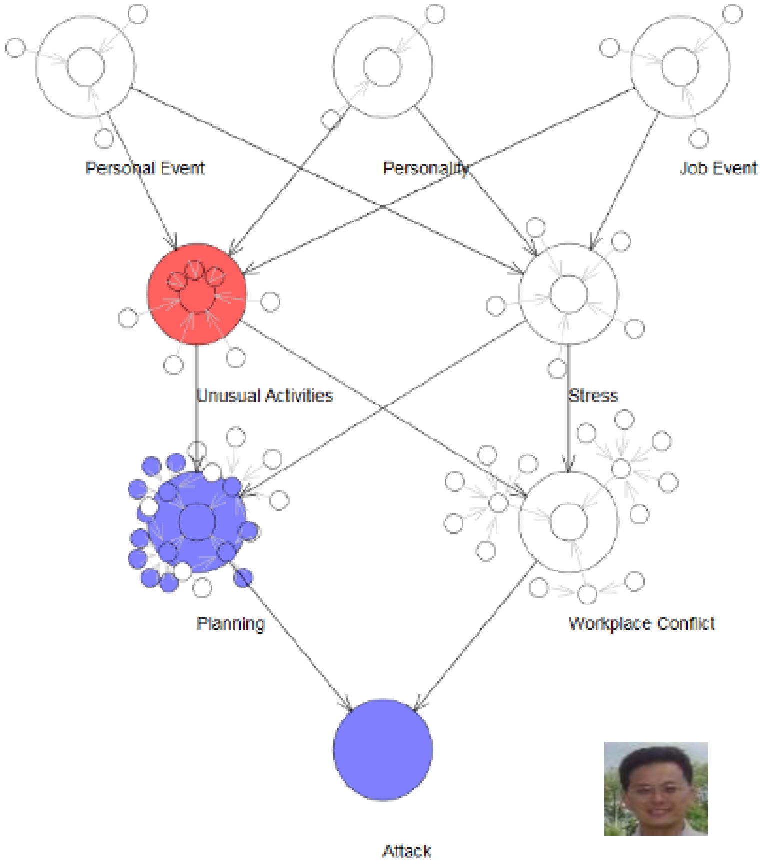
Complex Relationships

Topology and Flow



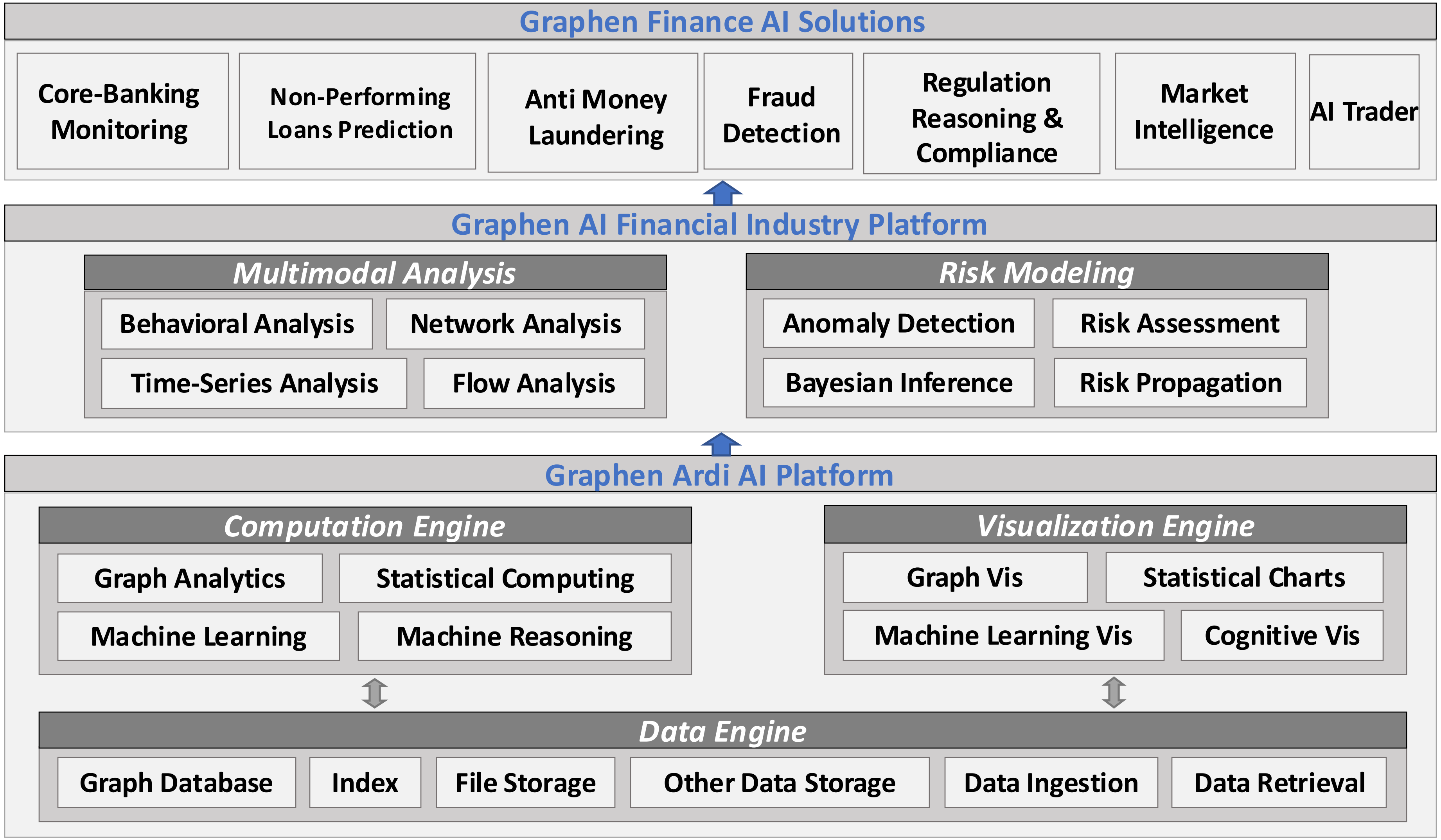
Reasoning

Causality and Prediction



Based on 15+ years of AI Security and AI Finance: usually accuracy is at least 2x-3x by context / relation analysis, 10x+ by reasoning / behavior prediction

Example: Graphen Financial Industry Platform and Solutions



Application Summary

Columbia University

- **Products:** AI offerings from the foundational platform to industry solutions
- **Main Business Models:** selling valuable industry solutions to business customers or jointly selling solutions with key business partners to users

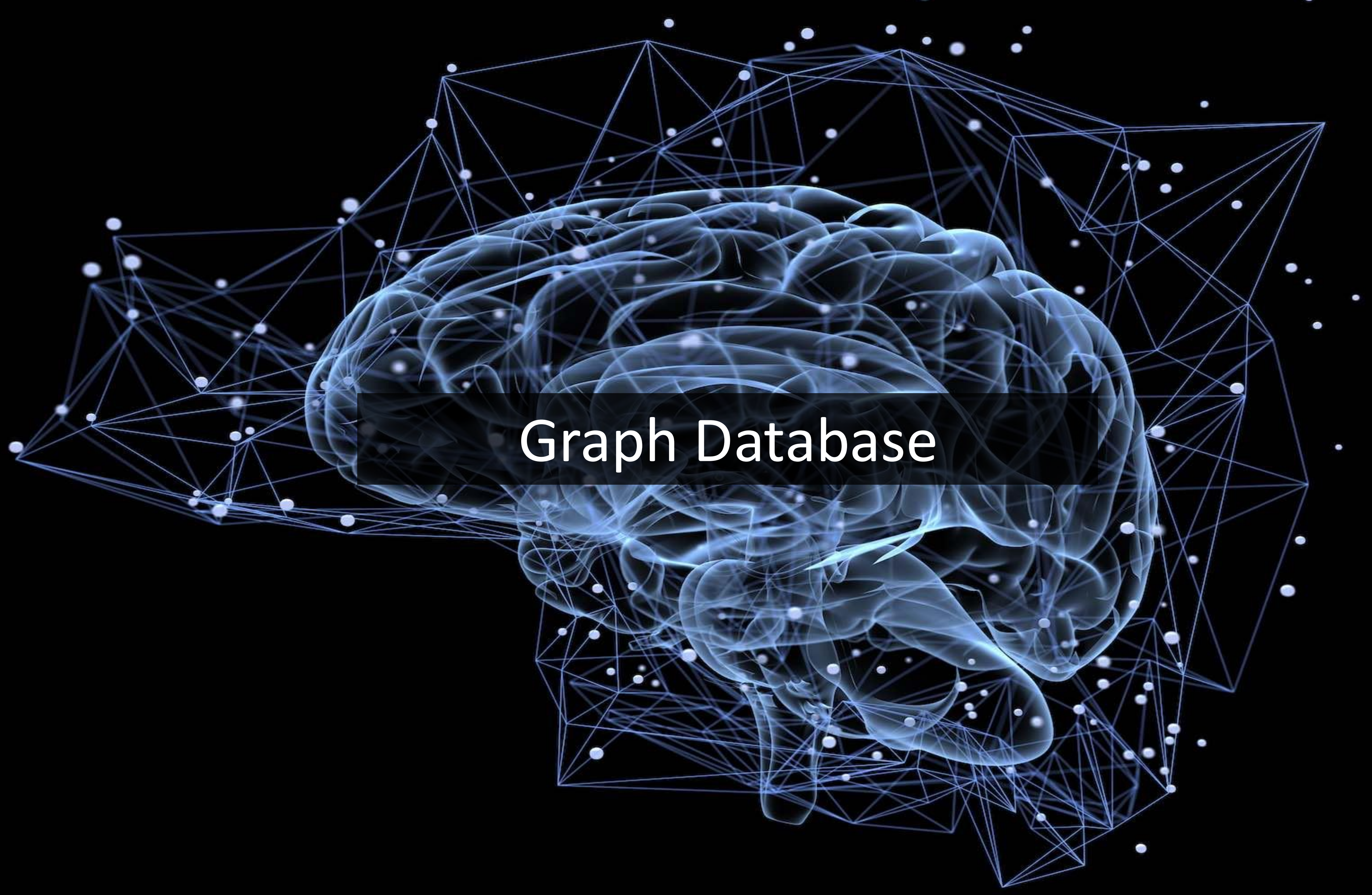

AI Foundation | Full-Brain Platform


AI Finance | Risk, Fraud, ESG & Intelligence


AI Medicine | Knowledge, Drug & Precision

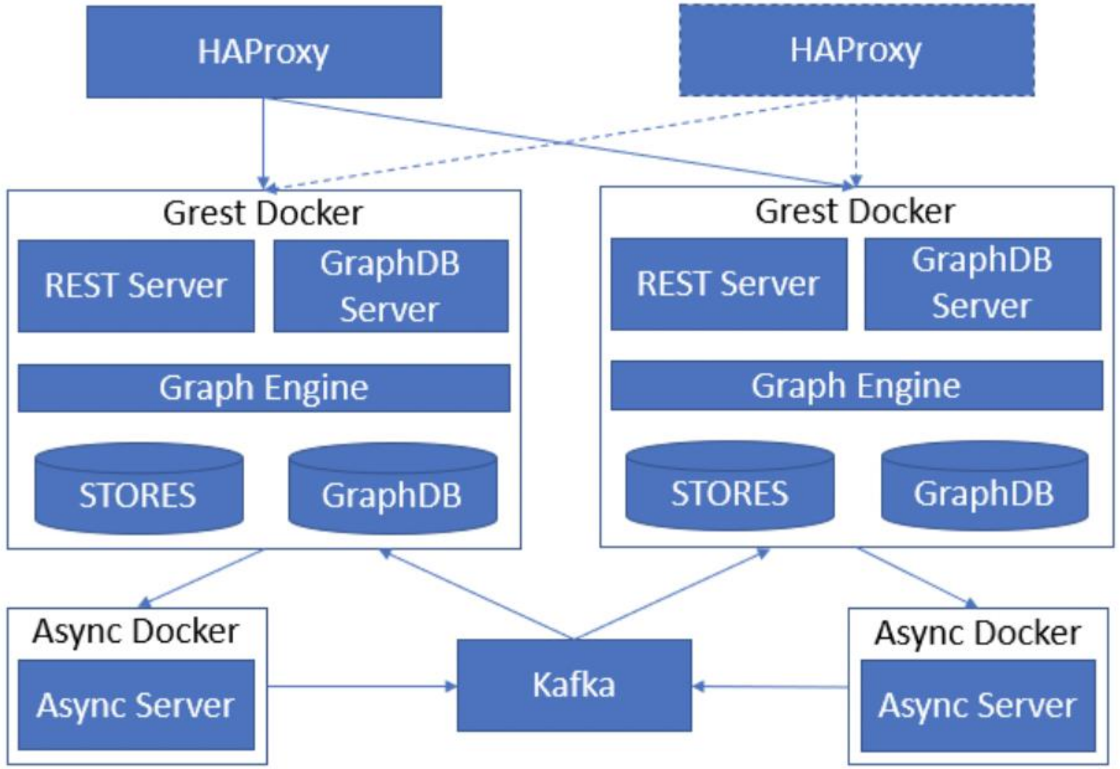

AI Automobile | Car Doctor


AI Energy | Clean Energy & Smart Grid



Graph Database

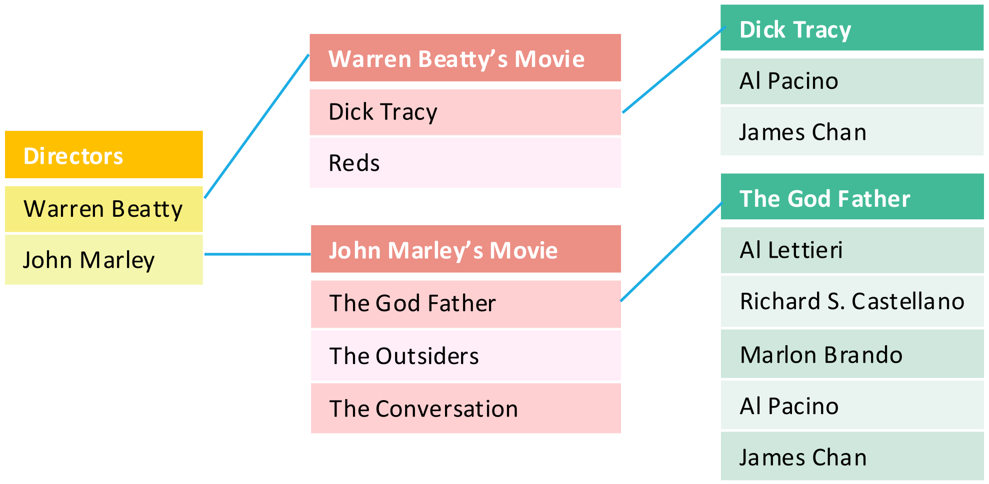
Ardi Graph Database is a C++-based native DB



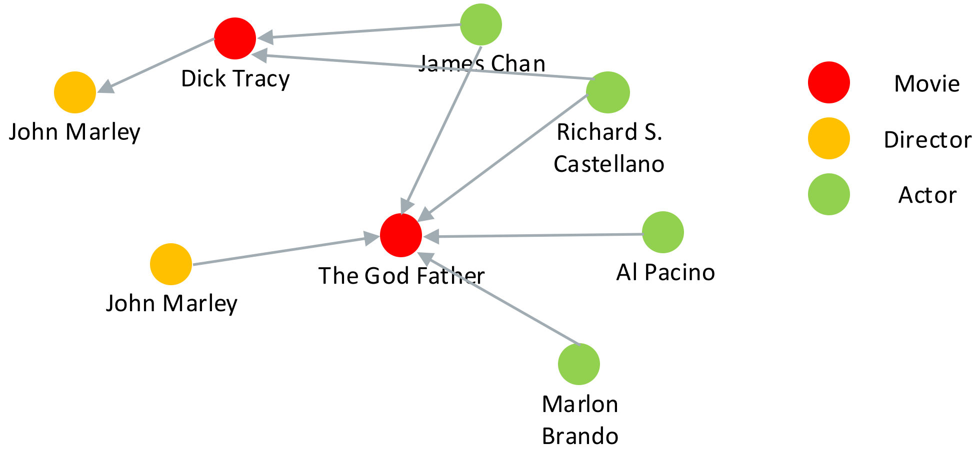
- **The storage layer** contains a high-performance **native graph store based on C++, indexers** to facilitate property-based search, **process store** to keep record of input/output associated with graph analytics and machine learning, and a **storage engine** to manage the graph database, indexes, and data store.
- **The analytics layer** has several computation modules utilizing Ardi platform's graph analytical algorithms and machine learning algorithms.
- **The query layer** includes all the services needed to synchronously or asynchronously load data into the storage layer, invoke graph analytics and machine learning algorithms, and retrieve data and results from running graph analytics and machine learning algorithms.
- **The visualization layer** enables raw graph data or computation results to be retrieved and displayed at the user interface via interactive visualizations.
- **The High Availability proxy servers** guarantees continuous system operations. Communications between layers are achieved via standardized APIs.

Graph Database vs. Relational Database

- SQL requires expensive join operations to compute neighborhoods of a vertex. Often the number of joins required is proportional with the distance from the source vertex as required by a specific algorithm.



Finding neighbors requires joins operations which may become intractable depending on the size of the database



Finding a neighborhood is logarithmic in the size of the database, thus tractable for any depth required

=> Ardi Platform supports both proprietary native Graph DB (by C++) and open-source relational database

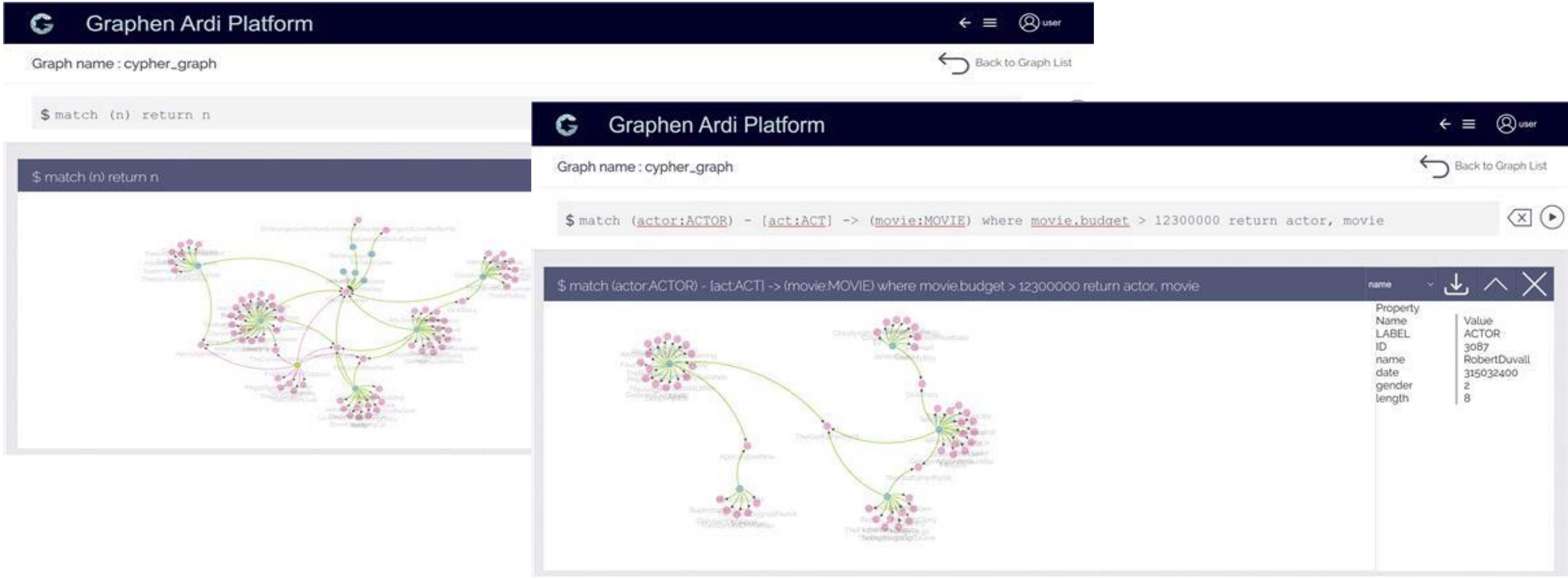
Graphen Database is Enterprise Production-Ready

Deployed in production in several largest banks in the world (in New York, Honk Kong, Shanghai and Taipei by Dec 2020):

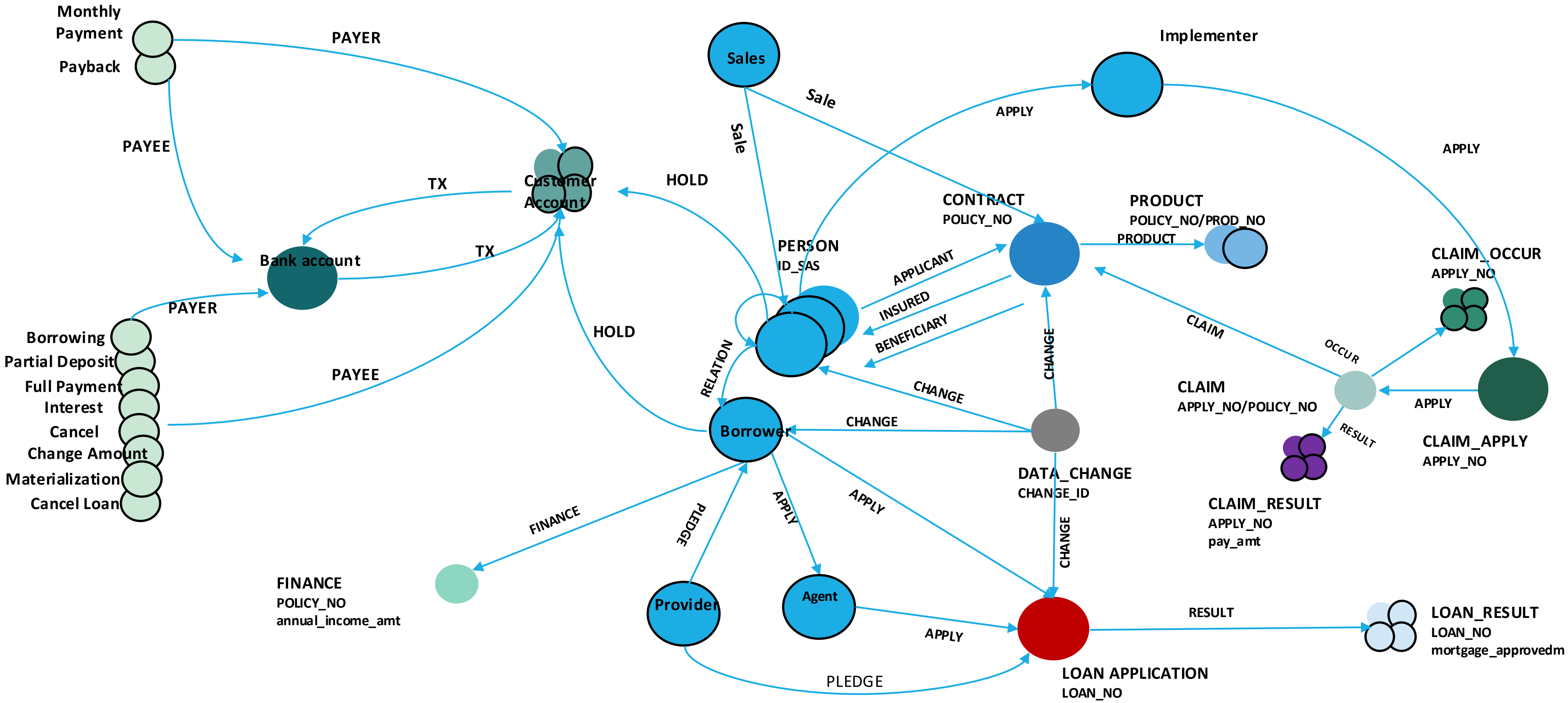
- Terabyte-sized native GraphDB, supports trillion of vertices and edges
- ACID-compliant and distributed Graph database and analytics
- Asynchronous job scheduling (both Autonomous ML and GraphDB)
- Scalable, distributed Analytics, modular and expandable through plugins
- Cluster, Replication and High-Availability with disaster recovery
- Error and event Logging, Monitoring, Backup and Recovery
- Supports both Graph Database and Relational Database
- Supports OpenCypher query language

Ardi Database's OpenCypher query support

Ardi Database supports OpenCypher query, making it easy to incorporate graph processing capabilities.



Fraud



Monthly Payment
Payback
PAYER
PAYEE
TX
Customer Account
TX
Bank account
PAYER
Borrowing
Partial Deposit
Full Payment
Interest
Cancel
Change Amount
Materialization
Cancel Loan
PAYEE
HOLD
HOLD
Sales
Sale
Sale
Implementer
APPLY
APPLY
CONTRACT
POLICY_NO
PRODUCT
POLICY_NO/PROD_NO
PRODUCT
PERSON
ID_SAS
APPLICANT
INSURED
BENEFICIARY
RELATION
CHANGE
CHANGE
CHANGE
CLAIM
OCCUR
CLAIM_OCCUR
APPLY_NO
CLAIM
APPLY_NO/POLICY_NO
APPLY
CLAIM_APPLY
APPLY_NO
RESULT
CLAIM_RESULT
APPLY_NO
pay_amt
Borrower
CHANGE
DATA_CHANGE
CHANGE_ID
FINANCE
PLEDGE
APPLY
APPLY
FINANCE
POLICY_NO
annual_income_amt
Provider
Agent
APPLY
PLEDGE
LOAN APPLICATION
LOAN_NO
RESULT
LOAN_RESULT
LOAN_NO
mortgage_approvedm

Ardi Graph Database Comparison

Performance Experiments on most graph functions common to Ardi, TigerGraph, and Neo4j

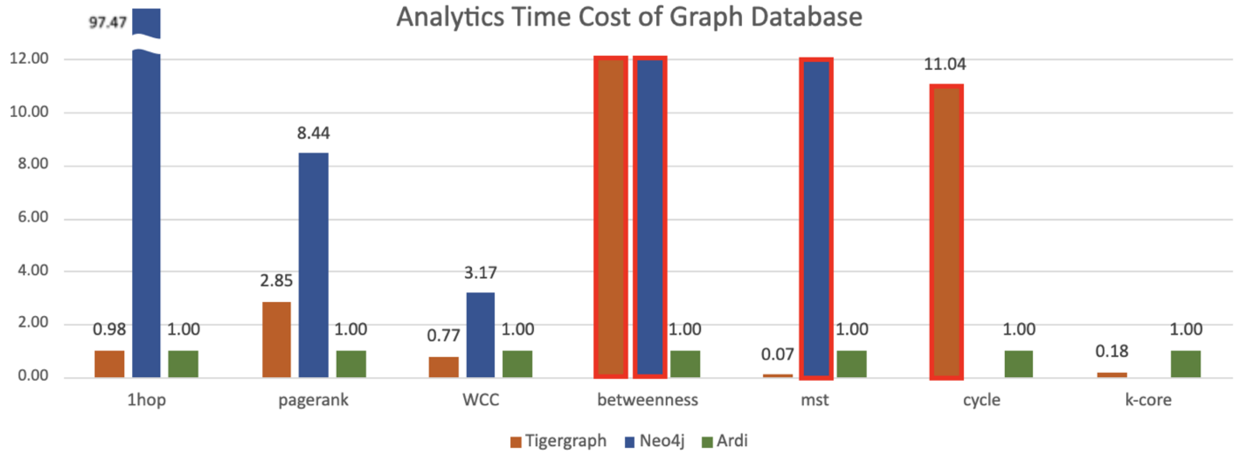
Performance experiment is conducted under following condition

Graph contains
- 2.4M vertices
- 67M edges

Testing Environment
- 32 cores CPU
- 128GB RAM

	Ardi	Tigergraph	Neo4j	Tigergraph/Ardi	Neo4j/Ardi
1hop	0.0766	0.0748	7.466	0.98	97.47
PageRank	10.829	30.838	91.384	2.85	8.44
WCC	42.004	32.401	133.134	0.77	3.17
betweenness	101.482	>7200	>43200	>70.95	>425.69
MST	1163.563	76.655	>43200	0.07	>37.13
cycle	19.035	>210.152	N/A	>11.04	N/A
k-core	214.512	38.814	N/A	0.18	N/A

- Ardi database outperforms Neo4j, which possess highest market share of graph database in all cases, including 1-hop neighbor, PageRank, Betweenness Weakly-Connected-Component and Minimum-Spanning-Tree. Besides,
- Ardi is competitive with Tigergraph, who claims beating most other graph databases in computation speed. Ardi outperforms TigherGraph in PageRank, betweenness and cycle finding, while being slower in MST and k-core.
- Therefore, we are confident to say that Ardi database is quite competitive among graph databases in the market.



* Neo4j timed out without results after 43000 seconds, and Tigergraph resulted timeout after 7200 second execution on calculating betweenness

** Tigergraph resulted in memory fault after 210 seconds and did not finish the cycle finding yet.

*** Neo4j did not support cycle finding and k-core finding in its Graph Data Science (GDS) library

Ardi Graph Database Comparison

Supported Algorithm

	Ardi	Tigergraph	neo4j
egonet	O		
Centrality	[1] Betweenness [2] Closeness [3] eccentricity [4] peripheral vertex [5] is central vertex	[1] Betweenness [2] Closeness	[1] Betweenness [2] Closeness [3] Harmonic [4] Degree [5] Eigenvector [6] ArticleRank
Pagerank	O	O	O
K-Core	O	O	
(Weakly/ Strongly) Connected Com ponents	O	O (wcc had been removed in version 3)	O
Louvain commu nities	O	O	O
Triangle counts	O	O	O
Cycle detection	O	O	

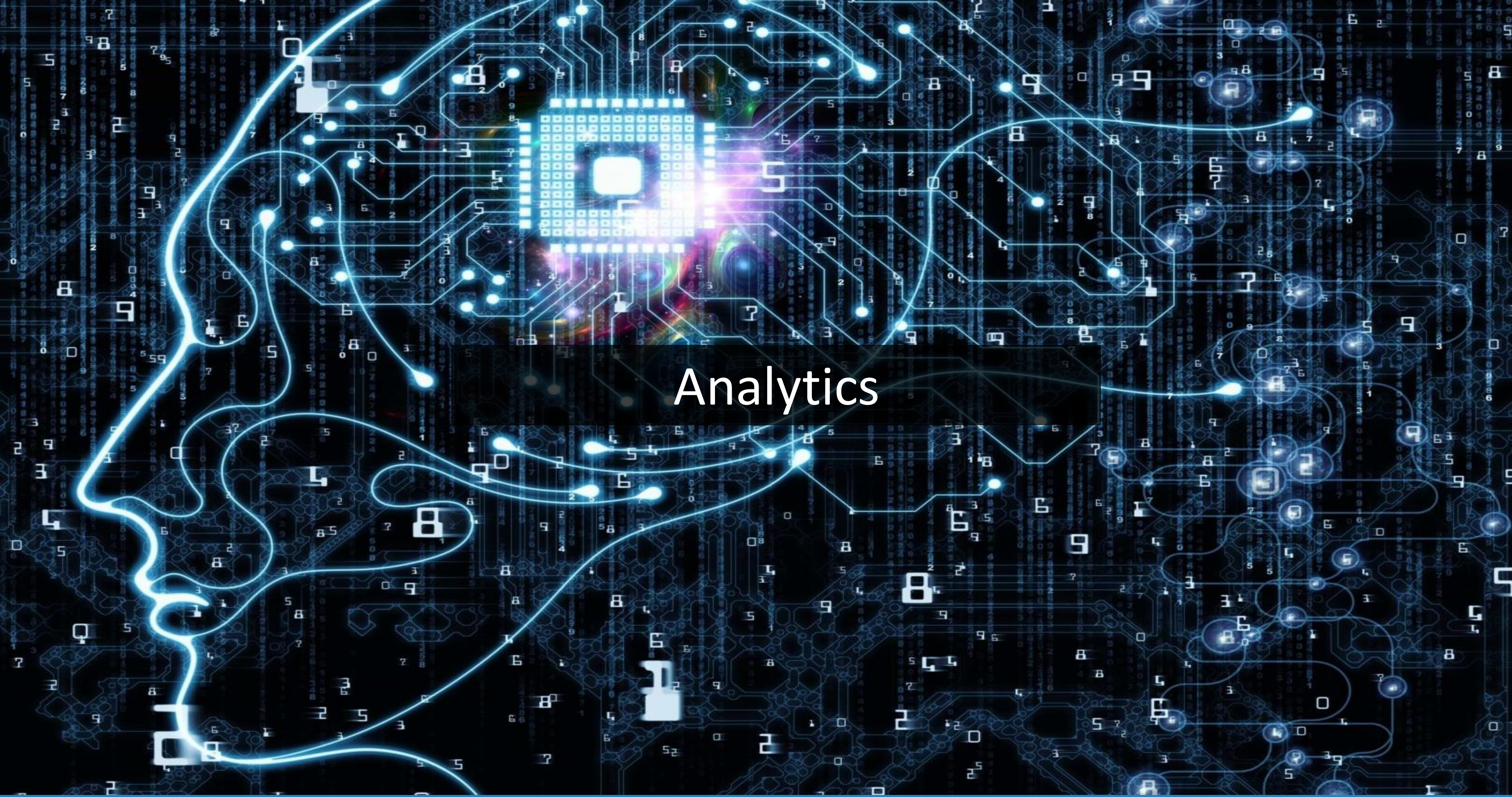
	Ardi	Tigergraph	neo4j
Cliques	O		
Local Clustering Coefficient	O		O
Maximal Independent Set		O	
Label Propagation			O
K-1 Coloring			O
Modularity Optimization			O
Heuristic estimate of graph diameter		O	
Spectral Clustering	O		
Entity Resolution	O	O	O

Ardi Graph Database Comparison

Supported Algorithm

	Ardi	Tigergraph	neo4j
Single-Source Shortest Paths	O	O	O
Yen's K-Shortest Paths			O
Shortest path between two vertices	O		O
All Pairs Shortest Path			O
A-star (an improved dijkstra algorithm)			O
Breadth First Search	O		O
Depth First Search			O
Random Walk			O
Minimum Spanning Tree (MST)	O	O	O
Minimum Spanning Forest (MSF)		O	

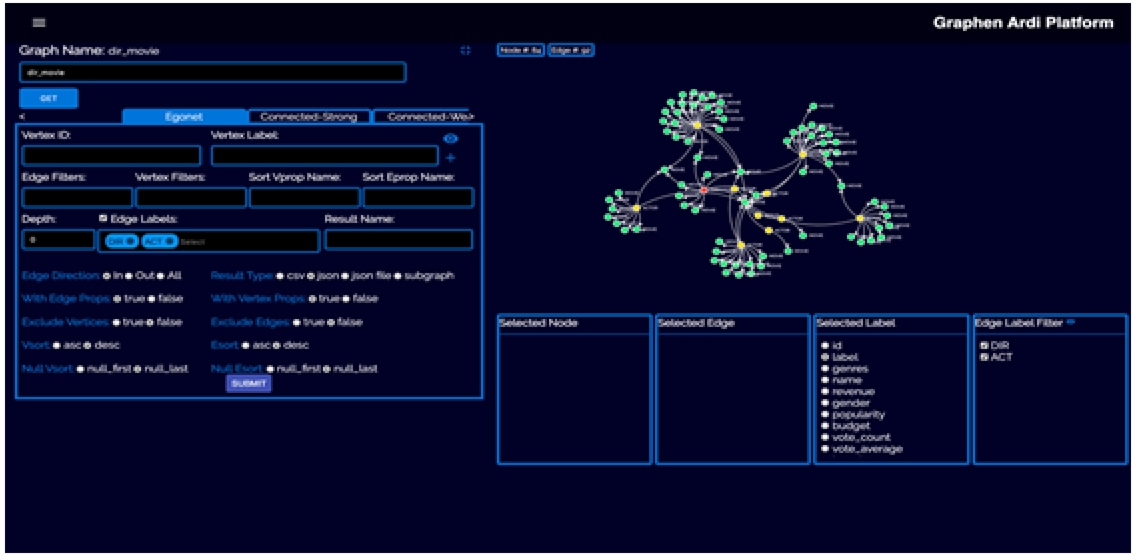
	Ardi	Tigergraph	neo4j
Link prediction	[1] number of common neighbors [2] jaccard similarity [3] jaccard similarity [4] salton index [5] leicht index [6] hub index [7] resource allocation index [8] number of all paths [9] shortest distance [10] katz distance [11] hitting time		[1] Adamic Adar [2] Common Neighbors [3] Preferential Attachment [4] Resource Allocation [5] Same Community [6] Total Neighbors
Similarity ranking	[1] Cosine Similarity	[1] Cosine Similarity [2] Jaccard Similarity	[1] Cosine Similarity [2] Jaccard Similarity [3] Pearson Similarity [4] Overlap Similarity [5] Euclidean Distance
k-Nearest Neighbor classification		O	O
Node embedding			[1] Node2Vec [2] GraphSAGE [3] Random Projection



Analytics

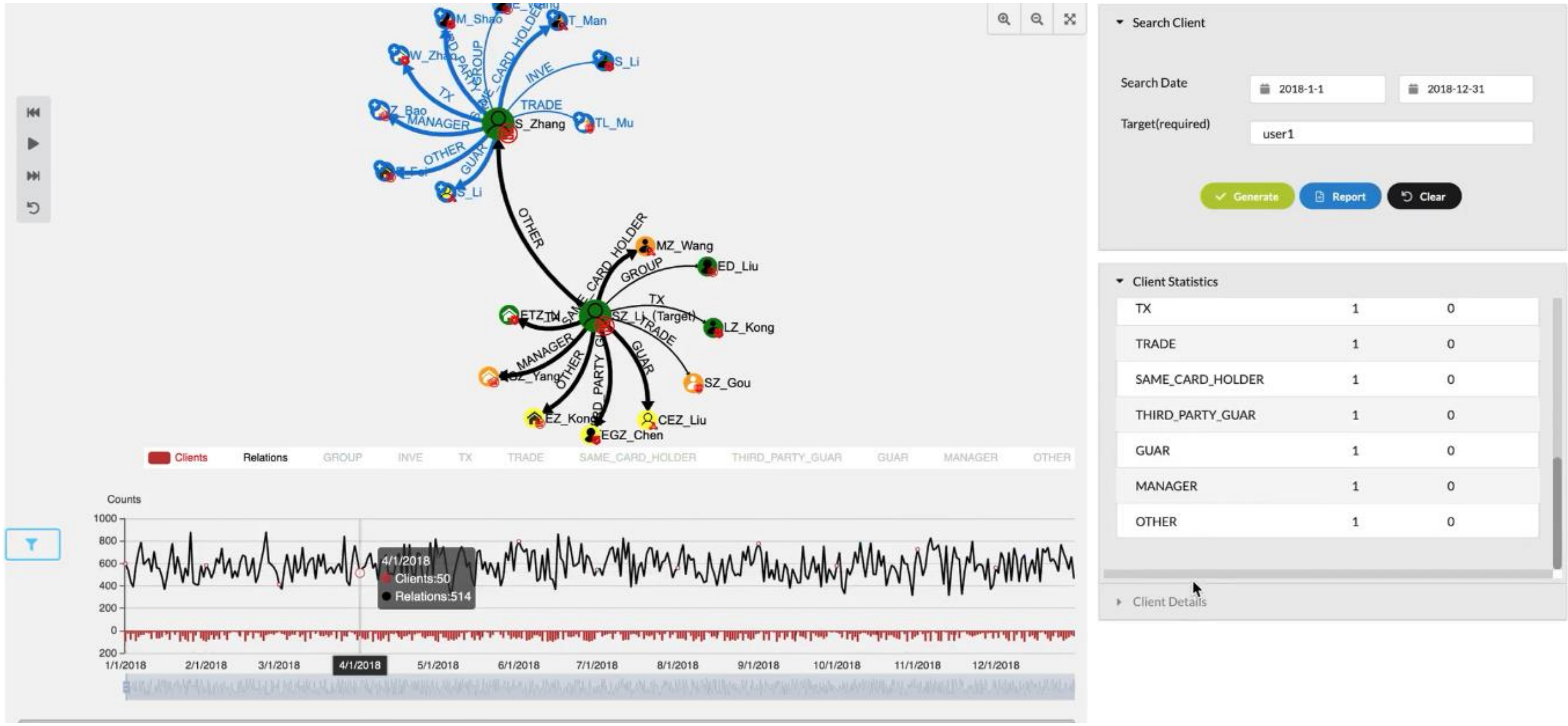
Ardi Graph Analytics Tools

- Support graphical analysis without any coding
- Efficient Analytics
  - Topological Analysis
    - Traverse
    - Shortest Paths
    - K-core
    - Minimum Spanning Tree
  - Metrics
    - Centrality and metrics
    - Compute PageRank
    - Link Prediction Indices
    - Clustering Coefficients
    - Similarity Ranking
  - Component Analysis and Retrieval
    - Cycles
    - Egonets
    - Strongly/Weakly Connected Components
    - Louvain Communities
    - Cliques
    - Graph Spectral Clustering
- Prediction
  - Missing Links Prediction
  - Entity Resolution
  - Risk Propagation



Example: Graph Analytics for Non-Performing Loan (NPL) Prediction

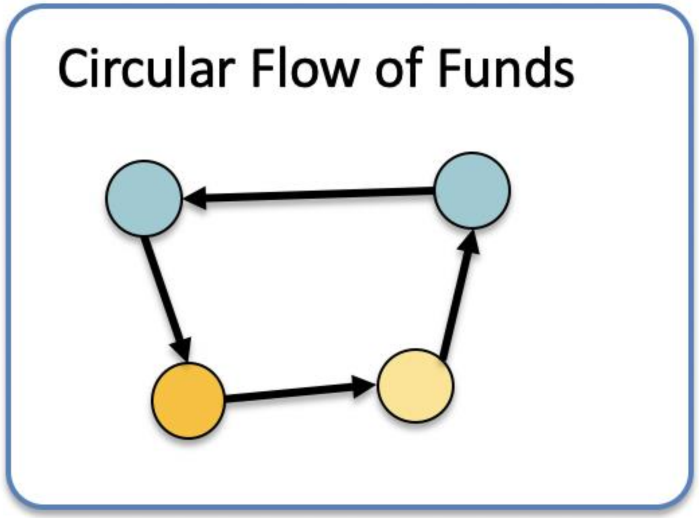
- Using Graphen Graph DB and Analytics, realized analysis of several millions of commercial customers. Processing time in customer's original system: **3 to 10 days**. Processing time using Graphen: **14 mins**.
- Using Graphen Cognitive Computing, Machine Learning, and Risk Propagation to realize NPL detection and prediction. Based on Graphen NPL Prediction to predict companies that might go default in the next month, the Average Precision of Top4 was **100.0%** and the Average Precision of Top10 was **92%**。



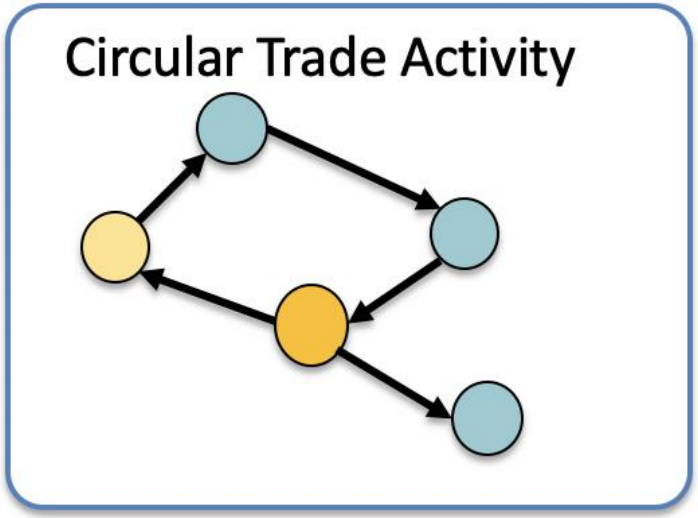
Relationship Generation for NPL

	Basic Elements	Implicit Relations
Investors & Shareholders	Investment • Majority Shareholder • Top Non-controlling Shareholder Executives • Corporate Representatives/Other Executives Other • Relatives/Spouse/Non-Spouse	• Same corporation legal representative • Legal Representative Outward Investment (Controlling/Non-Controlling) • Legal representatives with multiple jobs in different companies
Behavior Relations	Guarantee	General guarantee, Joint guarantee, Mutual guarantee, Guarantee Cycle
	Third-party asset collateral	Corporation legal third party
	Money Transaction	Transaction Cycle
	Trade • Factoring and Invoice Financing • Supply Chain Financing • Bank Notes • Letter of Credit • Entrusted Payment	Trade Cycle
United Credit Management of Group Customers		United Credit Management of Group Customers
Other Connection		• Inferred Relationship by Credit Card • Same Address/Phone number

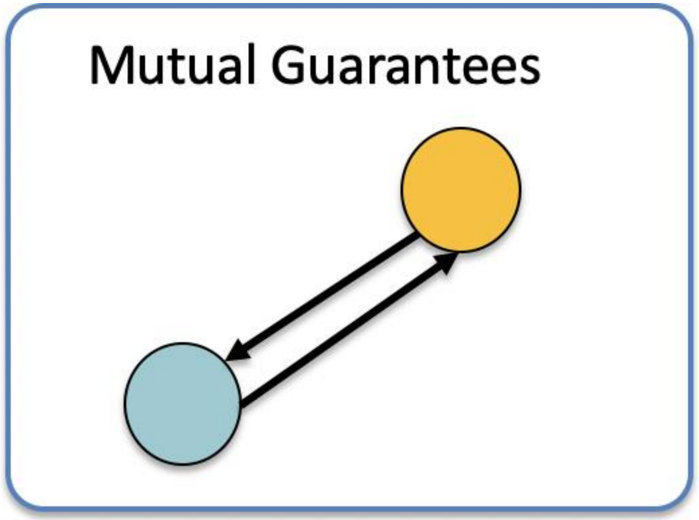
Critical Pattern Detection



The presence of a defaulting entity in the circular flow will increase the entity's risk



When the entity at the core of the relationship defaults, the risk of customers with large transaction amounts will increase

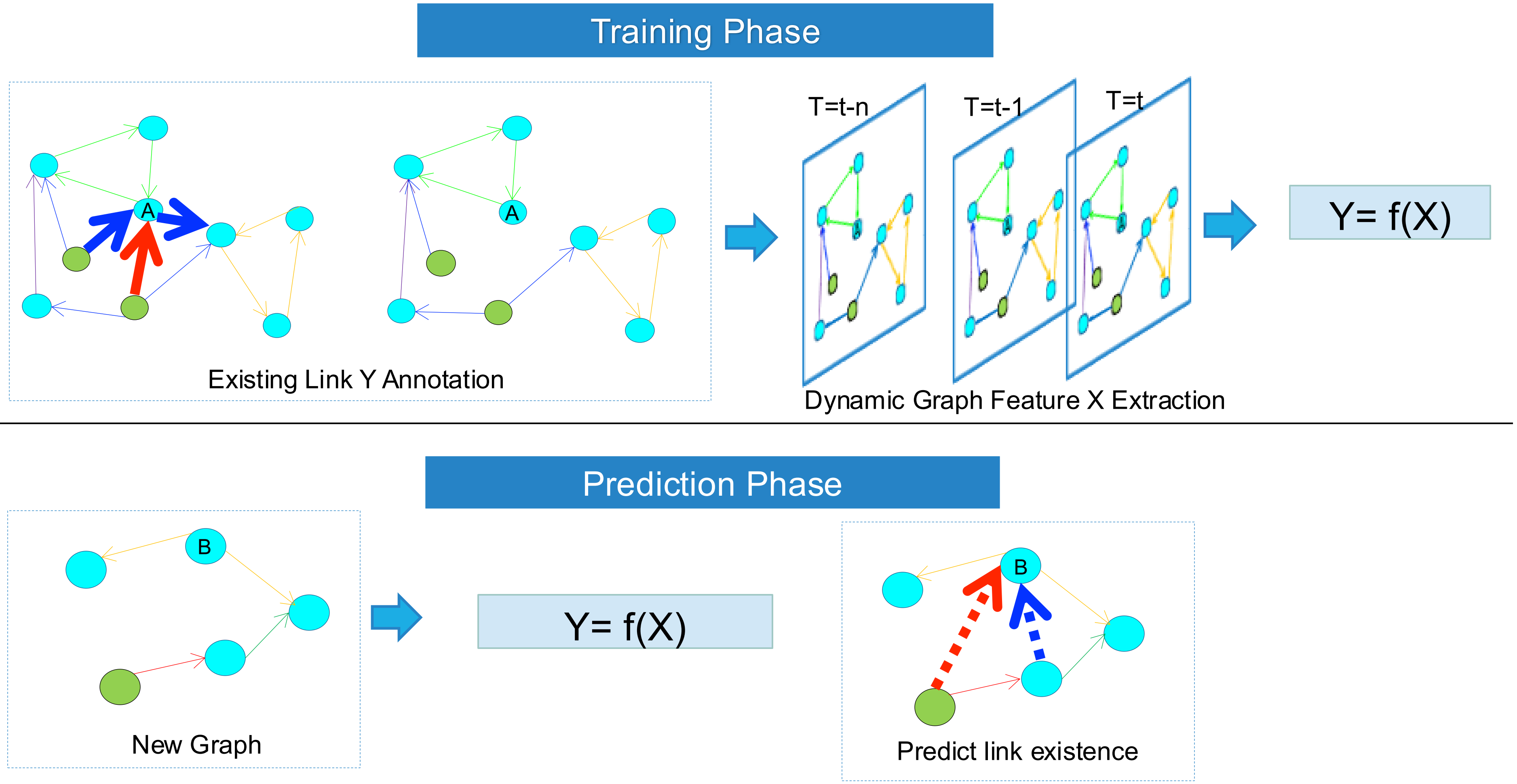


Mutual guarantees relationships are inherently suspicious

Critical Link Prediction

- Target: Given information on money transaction, trade and guarantee behavior pattern, Graphen's Link Prediction tool predicts the existence of potential connections between a pair of customers (or customer and corporation) among which at least one is loaned. We focus on the following relations.
  - Whether the customer is the spouse of the other
  - Whether the legal representative has invested in other companies
  - Whether the legal representative has another manager position in other companies.
- Using Ardi's Graph Analytics to extract graph topological features
  - Neighboring Coefficients: Common Neighbors, Jaccard's coefficient, Adamic/Adar Index
  - Distance Metrics: Weighted shortest path distance, Katz distance, Hitting Time
  - Cycle Analysis: Whether two vertices are in the same cycle
  - Missing Link Indices

Predicting Hidden Relationships



Detect potential missing links by predicting the probability of link existence between two nodes using supervised machine learning methods.

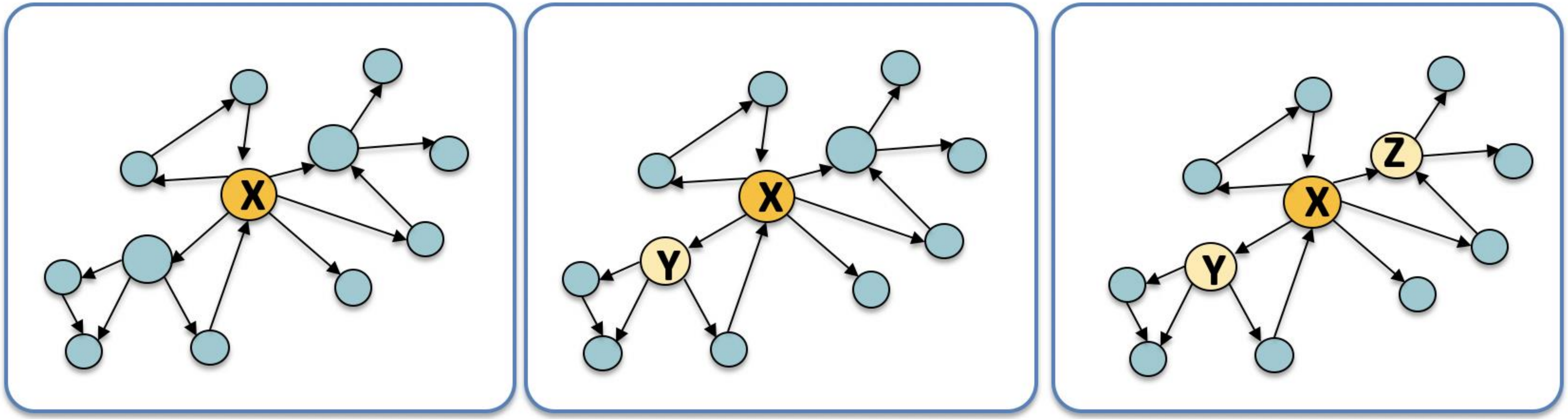
- Training Phase

  Model the relationship between X and Y using supervised machine learning algorithms

- Prediction Phase

  Calculate the link existence given the new graph and the relationship learnt in the training phase.

Risk Propagation

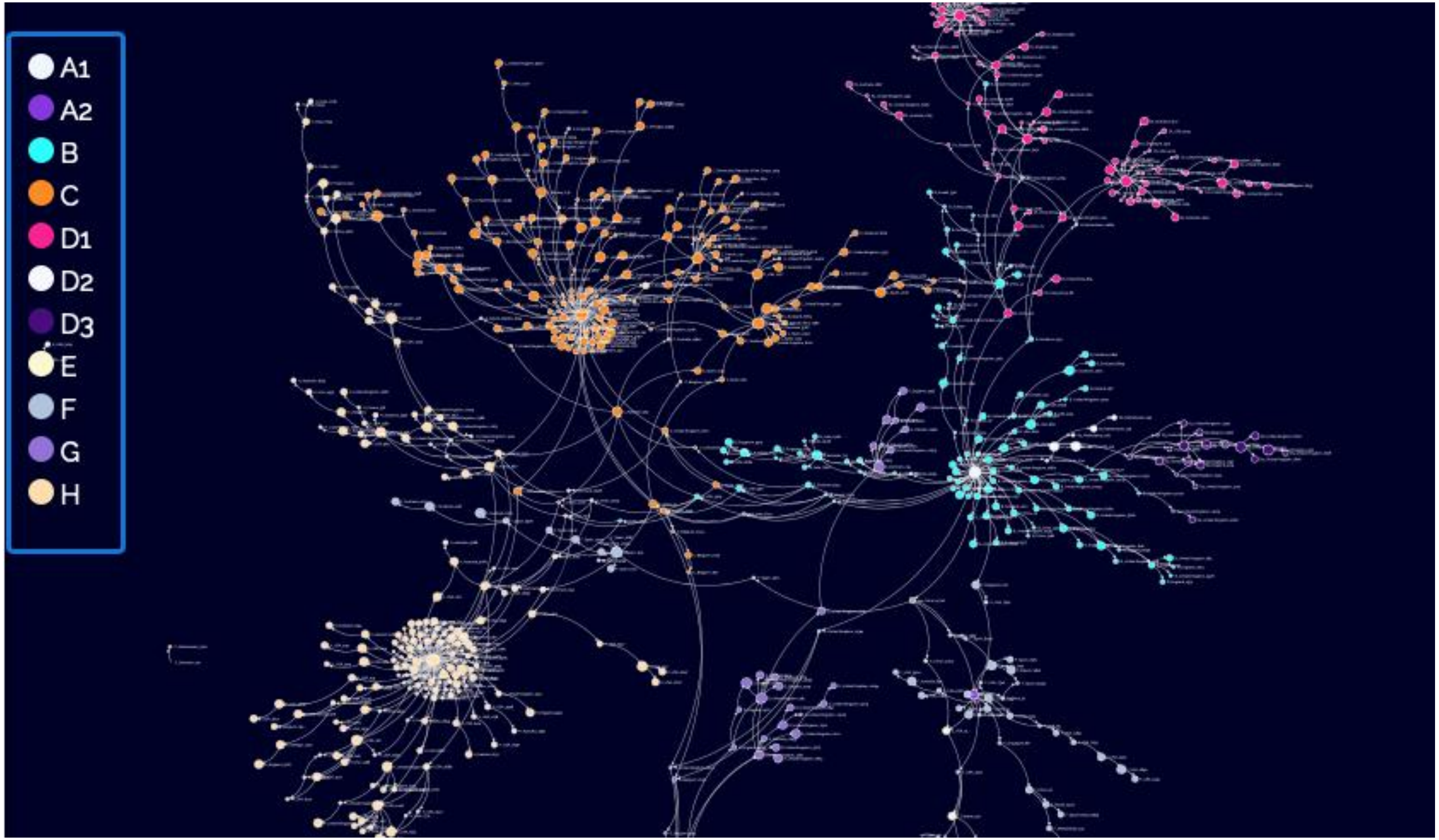


In time T1, a given entity X breaks contract

The solution predicts Y expected defaults in time T2

The solutions predicts Z expected defaults in time T3

Graph Spectral Clustering



- Multi-Resolution on Graph Topologies:
  - Finding key communities
  - Finding key bridges
  - Finding hubs
  - Finding visual analytics results that keep the original structure.
- Example: COVID-19 Worldwide Genome Evolution Graph

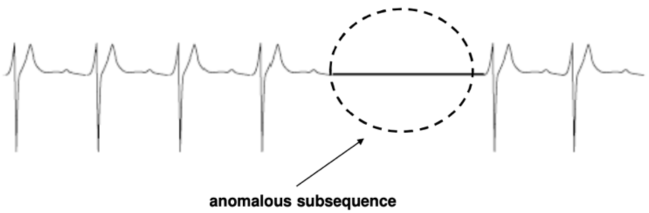
Anomaly Detection Tools

Unsupervised Abnormal nodes/links detection by estimating discrepancies with self or peer group behavior.

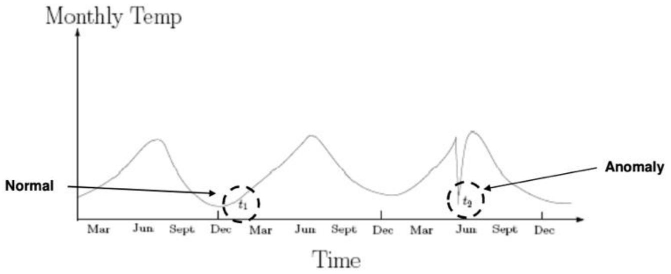
Technologies

- Statistical: Estimate a parametric model describing the distribution of the data;
- Proximity-based: Identify data points far away from the majority;
- Density-based: Identify data points in regions of low density;
- Clustering-based: Identify data points that do not belong strongly to any cluster.

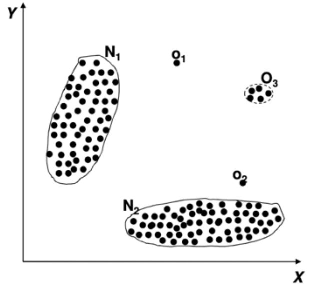
Types of Anomalies



Collective Anomalies



Contextual Anomalies



Point Anomalies

Example: Anti-Money Laundering

- Known entity risk
- Industry/occupation risk
- Geo-location risk
- Watch list hit
- Negative news hit

Entity Risk

Network Risk

- Related individual risk
- Related company risk
- Transacting individual risk
- Transacting company risk
- Network change

Risk Score

- Rejected transactions
- Alert/case/SAR history
- Rules triggered in alerts from existing systems
- Sanctions against organization/ location

Red Flags

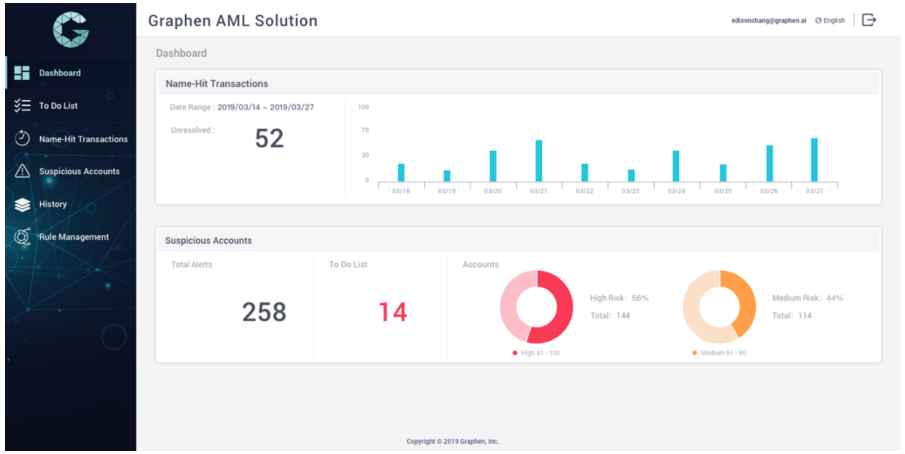
Machine Learning

- Abnormal Behavior: Deviate from historical data and peer behavior
- AML Model: Trained by historical dataset

1. Decreasing False Positive
2. Finding Unknown Unknowns

Deployment Examples:

- ~80,000 SARS *-> reduced 80%*
- ~100,000,000 Clients
- ~200,000,000 Accounts
- ~1,000,000,000 Relations
- ~200,000,000 Transaction



Example: Entity Risk

Industry

- Discrepancies between KYC data and external data

Geography

- AML risk identified with customer’s location
- Discrepancy between self-reported location and external data

Negative News

- Negative news related to the customer, their employer, associates, etc.

PEP

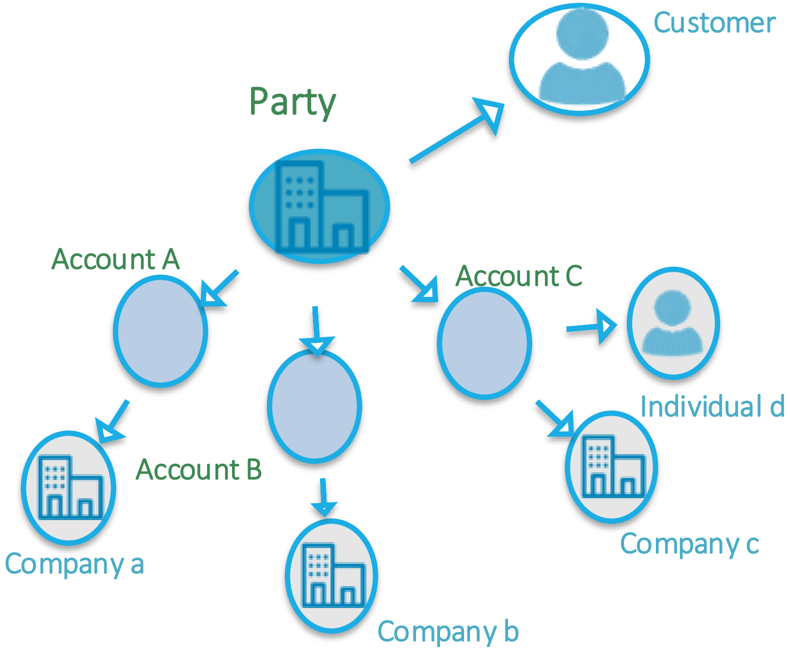
- The customer
- appears on a PEP list

Information Source

• Onboarding/KYC data • CDD、EDD data • Client location • Bank Branch location	• **Open Internet / social media** • **Online news sources** • **Third-party vendors**

Example: Network Risk

A graph depicting all entities connected to the "party" (the customer being analyzed)



- **Determining risk from the network graph:**

  The graph will be analyzed for certain features used to calculate a Network Risk Score.

Relationship Types

- Funds/Capital Relationships
- Transaction Relationships
- Business Relationships (employer/employee, contractors, etc.)
- Stock Ownership
- Other Relationships (marriage, etc.)

Party

Focus of analysis

Company Individual

Account

Ardi Applications - AI Finance

– A Full-Spectrum AI Finance Solution Provider

Central Monitoring

Real-Time monitoring the operation of entire bank(branches, ATMs, mobile banking, customer services, social media, etc.)

Anti-Money Laundering

Using AI to detect risking money laundering schemes.

Regulation Inference

Using ML, NLP, and Reasoning to effectively track and analyze regulations for better compliance.

Market Intelligence

Building Knowledge Graphs by gathering news, judging company's public ESG images, and predicting financial markets.

Cybersecurity

Protecting Financial Hubs with advanced behavior understanding system and intention prediction.

Non-Performing Loan Prediction

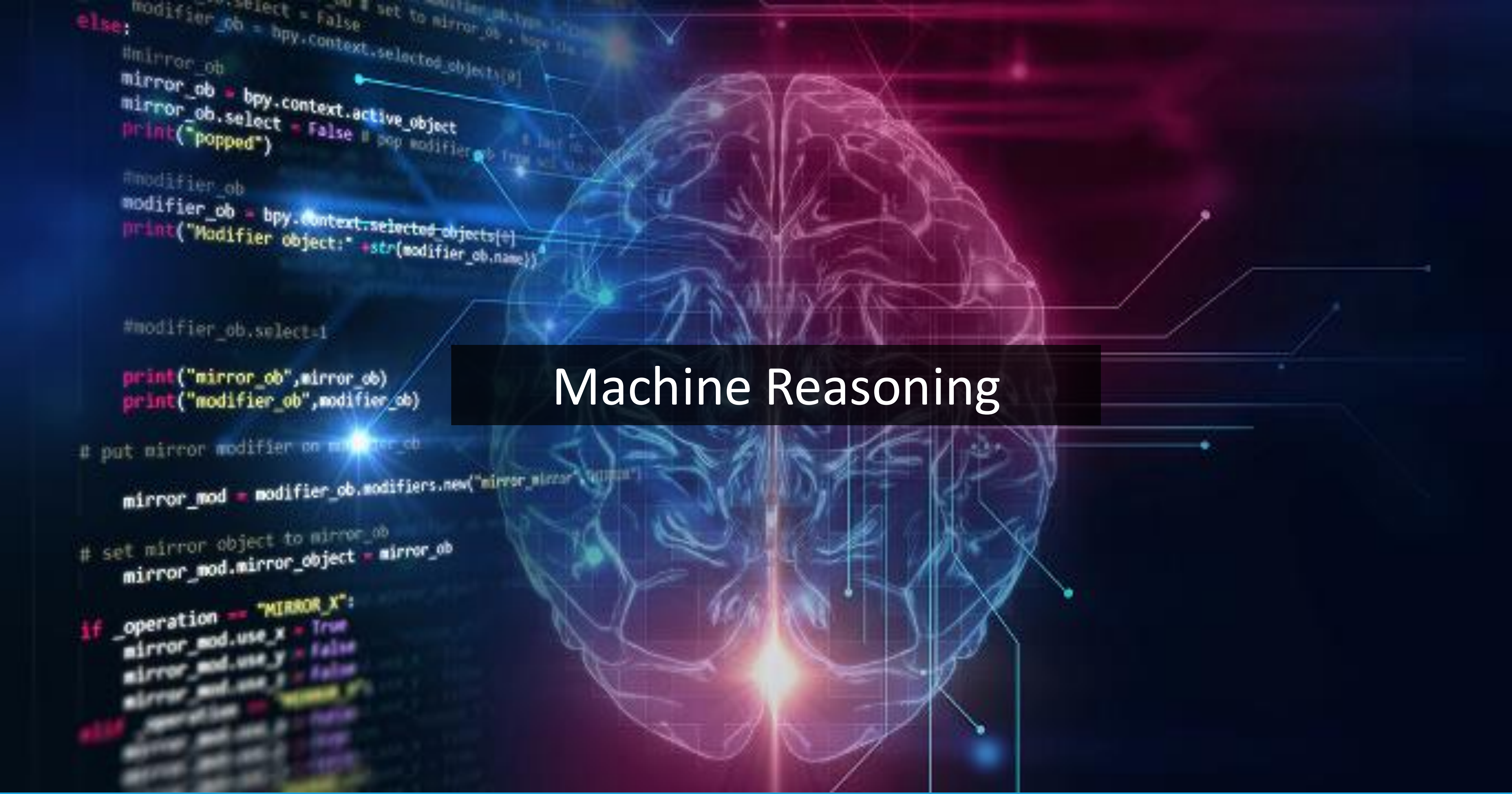
Analyzing relations of accounts and their risk propagations.

Fraud Detection

Using Advanced AI to automatically detect all kinds of fraud behaviors, including agents in banking & insurance industries, and customers.

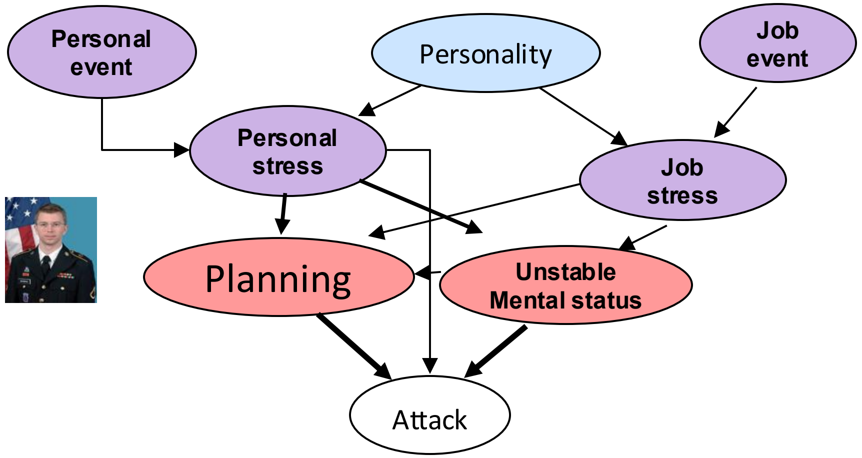
Trade Finance Due Diligence

Automatic process to crawl data and investigate trade entities.



Machine Reasoning

Ardi Machine Reasoning



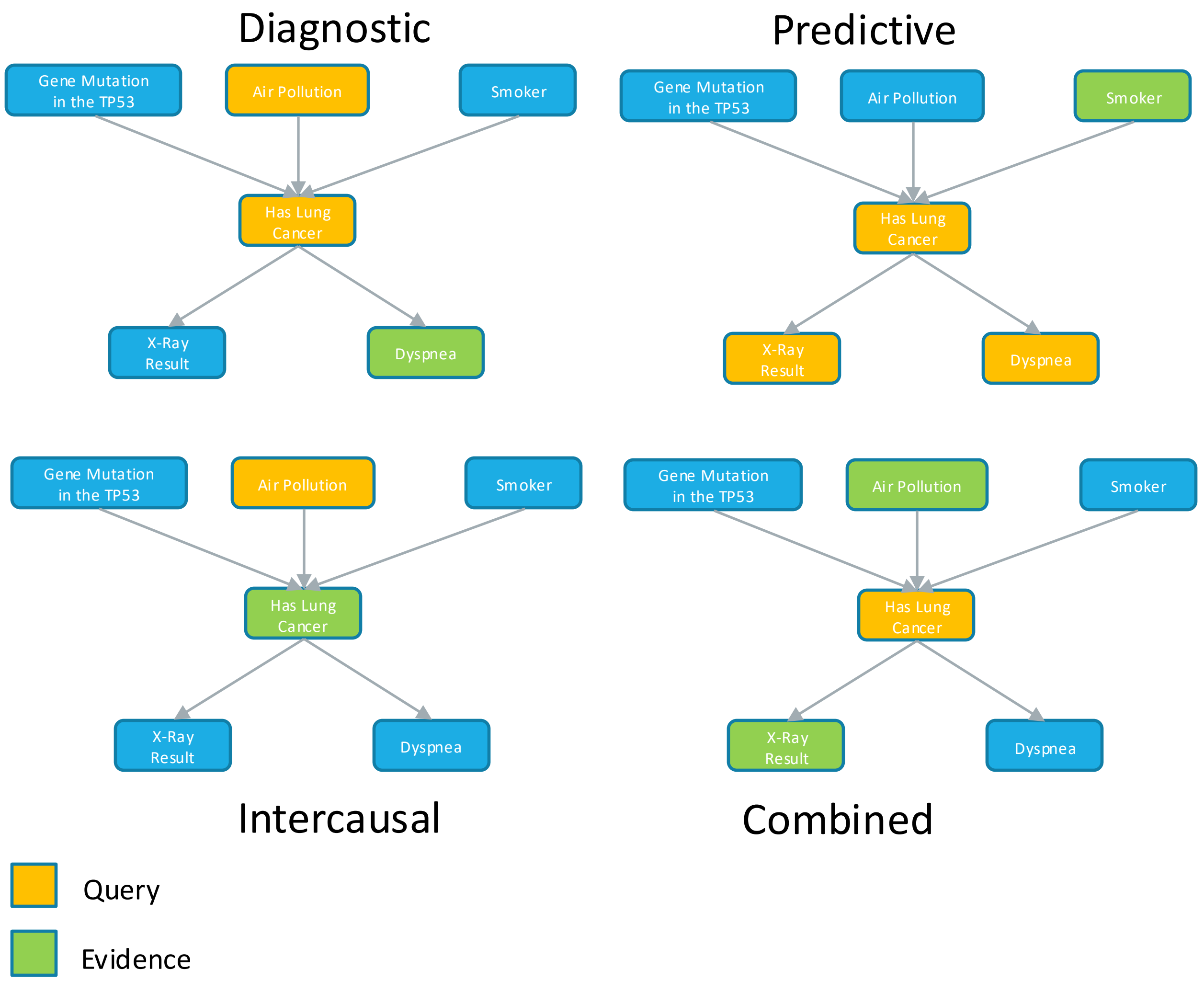
One of the main challenges in building an efficient system is the ability to learn and to reason under uncertainty, and one of the most successful approaches for dealing with this challenge is based on the framework of Bayesian Networks.

Bayesian Networks offer an expressive visual and quantitative tool for

- Learning and representing reasoning procedures
- Understanding causality among variables

- Machine Reasoning may improve risky behavior prediction accuracy up to 10x.

Types of Reasoning

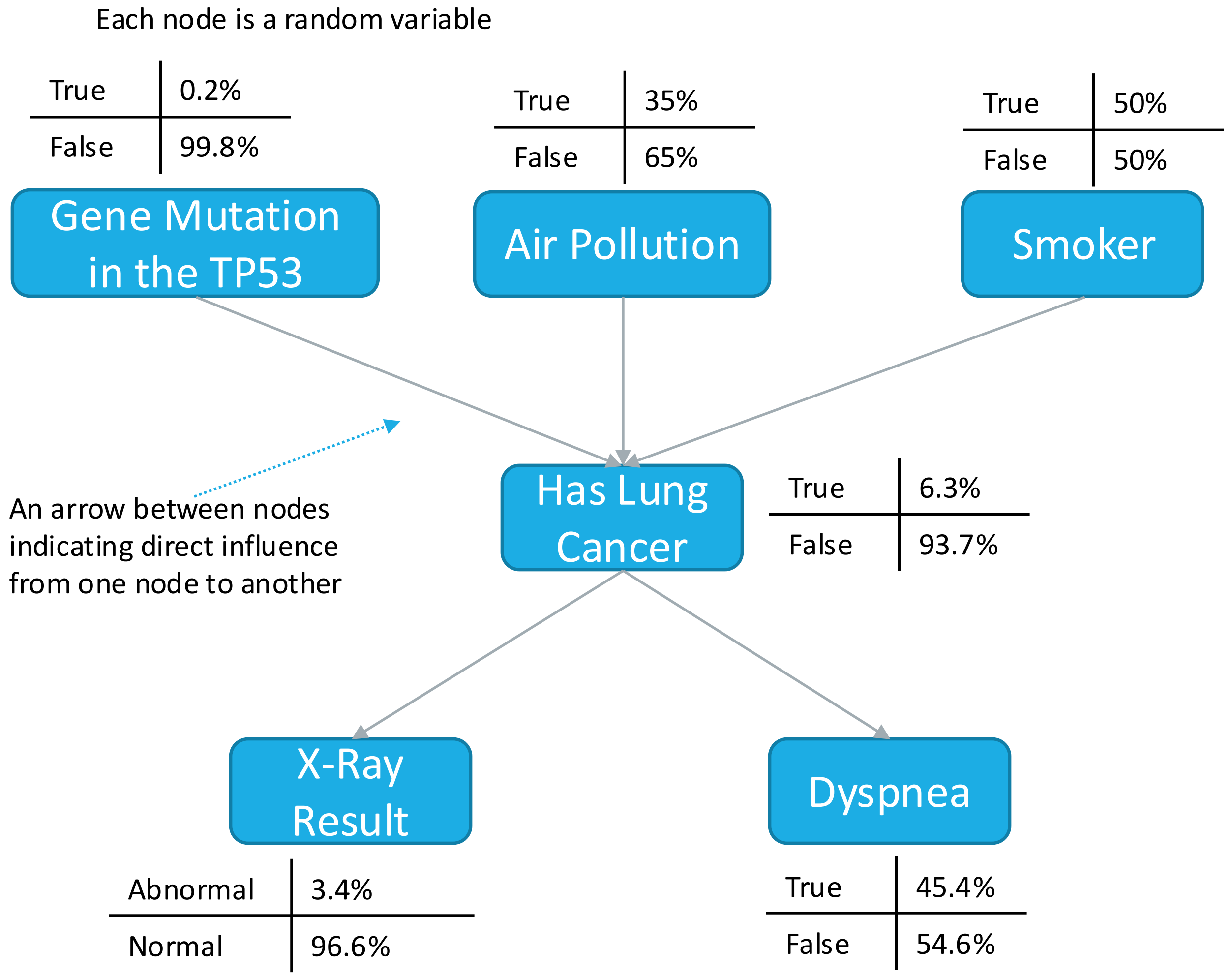


- **Diagnostic:** Given evidence about an effect, how does this change the belief in this causes?
- **Predictive:** Given evidence, what are the predicted outcome?
- **Intercausal:** Given evidence about a cause and about its effect, how does it change the beliefs in other causes?
- **Combined:** Given evidence about background causes and effects, what are the new beliefs in intermediate nodes?

Why using Bayesian Networks for Reasoning?

- Graph representation of real-world data
  - Conditional independencies & graphical language capture structure of many real-world distributions
  - Graph structure provides more insight into domain and allows in-depth domain knowledge discovery through network construction
  - Expert prior knowledge may often be incorporated when learning the graph structure
- Learned Bayesian model solves analytical limitations
  - Learned model can be used for many tasks
  - Supports all the features of probabilistic learning
  - Deal with missing data & hidden variables

Bayesian Networks



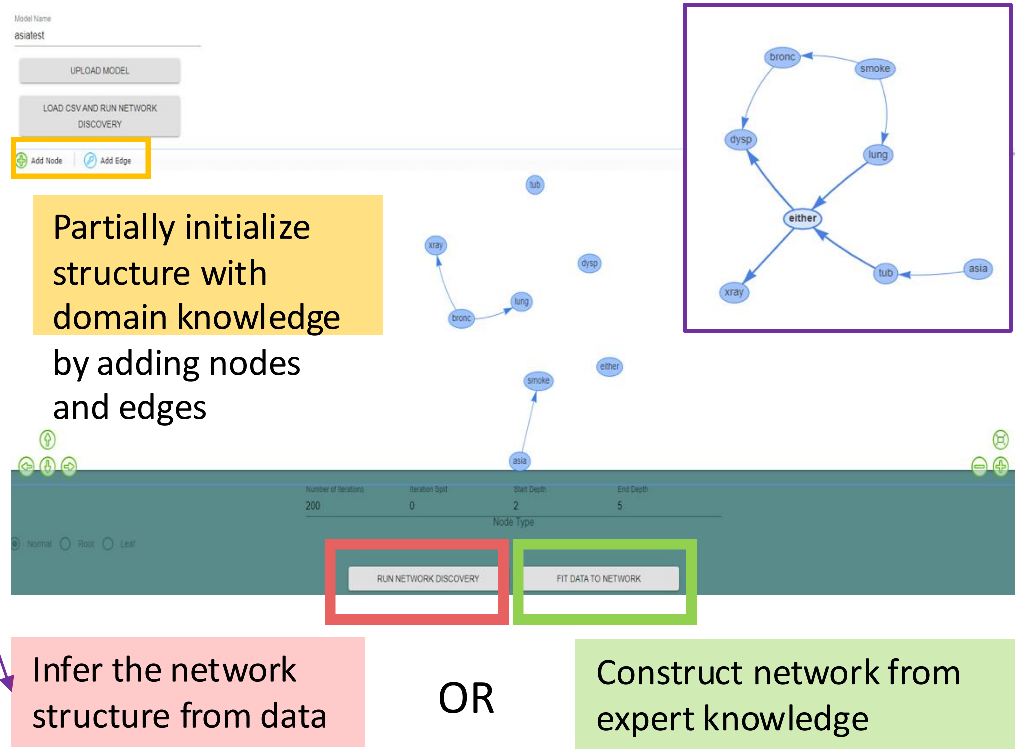
- A network model that follows the structure of a directed acyclic graph (DAG), G=(V,E), where V denotes nodes and E denotes edges;
- Encode the conditional independencies of each vertex given its parent, measuring how the change of one variable affect others at different levels;
- A Generative model that allows arbitrary queries to be answered.

Reasoning Structure Inference

- Target
  Given a set of random variables, find the optimal Bayesian network with best structure and parameters that captures the casual relations between variables.

- Score-based Model Selection Criteria
  - Cooper-Herskovits (CH) Criterion
  - Bayesian Information Criterion (BIC)
  - Minimum Description Length (MDL)
  - Akaike Information Criterion (AIC)

- Model Optimization
  - K2 search for model with highest CH Criterion
  - Random restart hill-climbing
  - Tabu Search



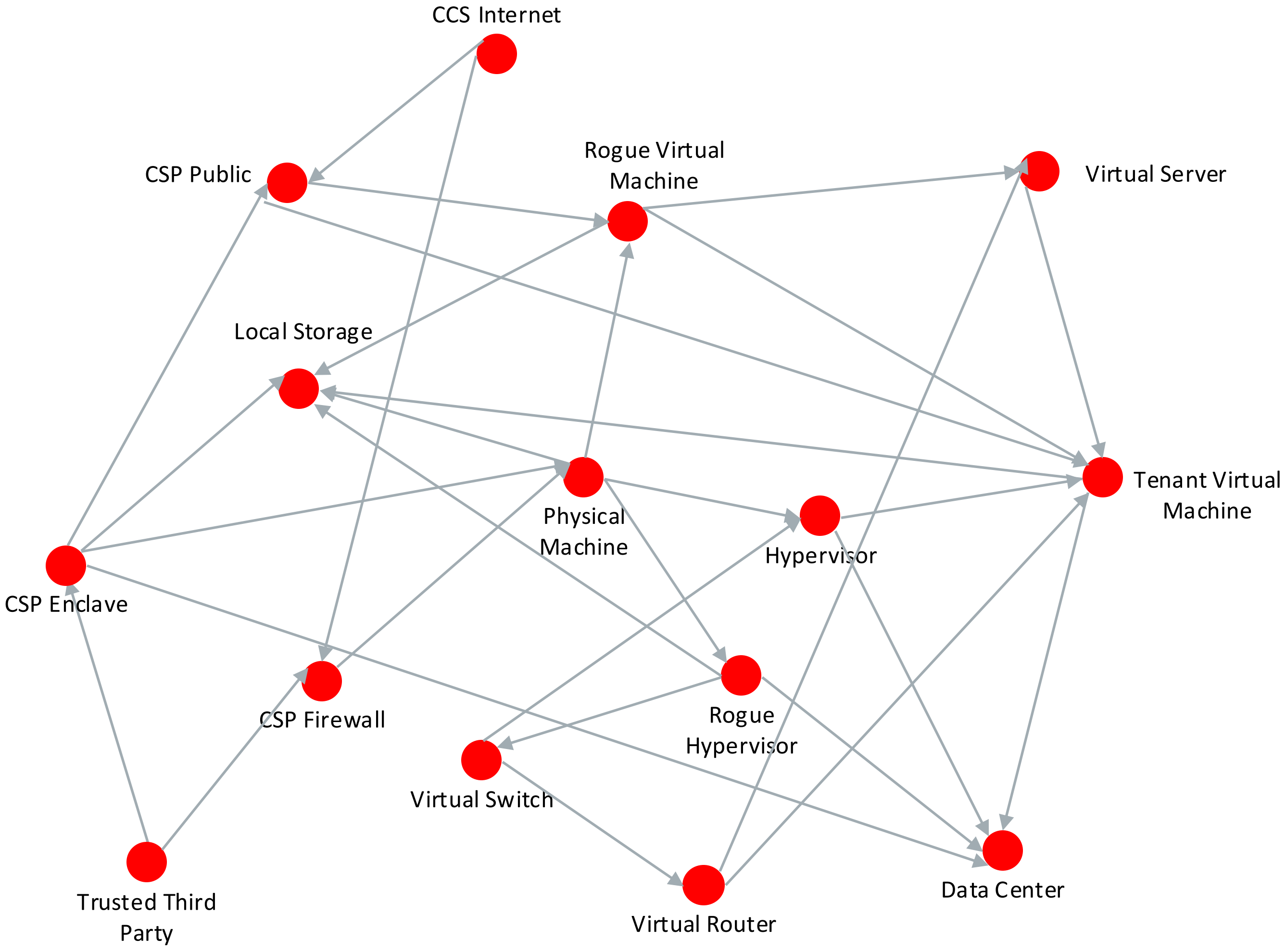
Graphen Ardi Bayesian Network GUI

Example: Bayesian Inference in Cyber Security

- APT attackers possess high levels of technical skills and have extensive resources at their disposal, and this has enabled them to effectuate sophisticated stealthy reconnaissance, surveillance and data exfiltration attacks with little traceability if any at all. The threat actor executes a series of coordinated actions to obtain a set of assets needed to reach the goal(s).

- Target:
  - Predict potential attack in the next stage
  - Evaluate the likelihood of an APT occurring
  - Model the uncertain aspects in cyber security
    1. The uncertainty on attack success
    2. The uncertainty of attacker choice
    3. The uncertainty from imperfect IDS sensors

Bayesian Inference on Cyber Anomalies



Construct attack graph from domain knowledge

Conditional Probability Assignments

Inference
Query the probability of the vulnerability when a new event occur

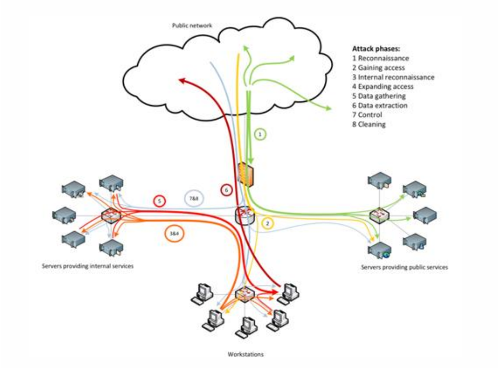
Detection Long-Term Anomalies and Threats

Stage 1

External Reconnaissance

- Uncharacteristic site usage
- Threat intelligence feeds
- Dark Web forum chatter

- Web Activity
- Email Activity



Stage 2

Gaining Access

- Attachments
- Difference in file sizes
- Common traits of phishing mails

- Exploits delivered through phishing mails or infected websites

Stage 3

Internal Reconnaissance

- Network usage changes
- Query Network information

- Using malware, botnet and rootkit to obtain local privileged access

Stage 4

Expanding Access

- Different traffic patterns
- Traffic content
- AD logs
- Network Activity

- Use malware and exploits to achieve domain-level accounts

Detection Long-Term Anomalies and Threats

Stage 5

Gathering Information

- Use Malware or botnet to identify directory storage

- File usage pattern changes
- Services use changes
- Quantity of files accessed increases

Stage 6

Information Extraction

- Data is encrypted and extracted

- Volume of traffic
- Network Activity

Stage 7

Control of Information leaks

- Fraudulent transaction is executed

- Internet usage (segments, sites)
- Traffic contents
- Endpoint logs
- Network Activity

Stage 8

Erasing Tracks

- Data, logs are removed

- Changes to system
- Changes in file sizes
- Gaps in log files
- Usage of erasing tools

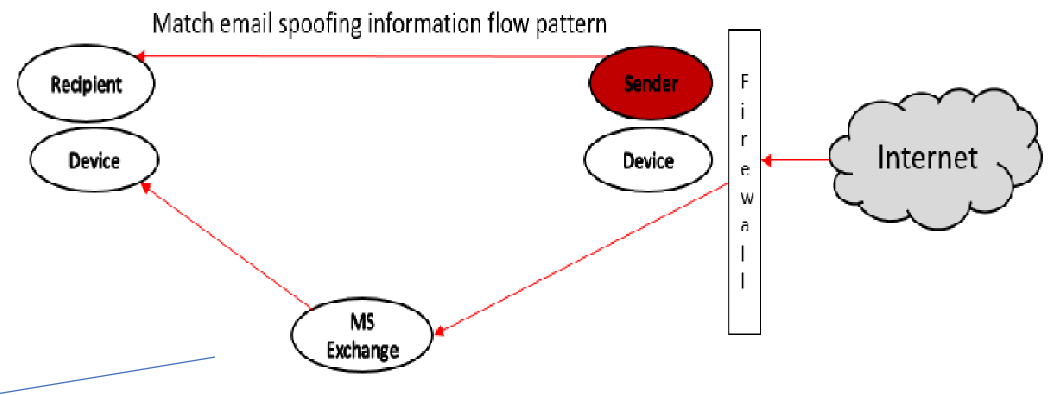
Machine Reasoning to Aggregate Risk

Columbia University

Behavior-based anomaly identified

- **Historical/group behavior: a small number of emails to a few frequent contacts per day during business hours**
- **Current behavior of employee A: an email sent outside normal business hours to all his contacts**

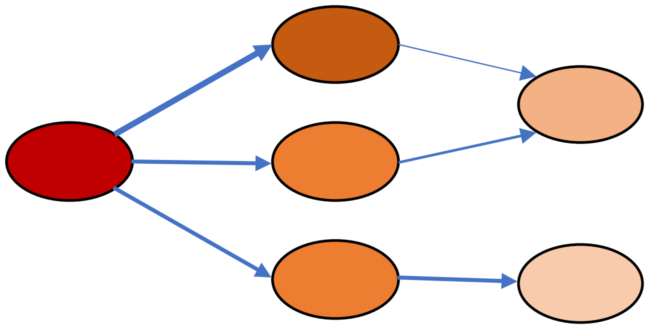
Flow-based anomaly identified



An email that appears to be from employee A sent to all his contacts matches email spoofing information flow pattern

Risk Aggregation → Risk Propagation → Final Risk Assessment

A single risk score for each entity to indicate its overall risk



For the entity that has elevated risk, propagates risk to other entities connected to it

Graphen Automobile – AI Car Doctors

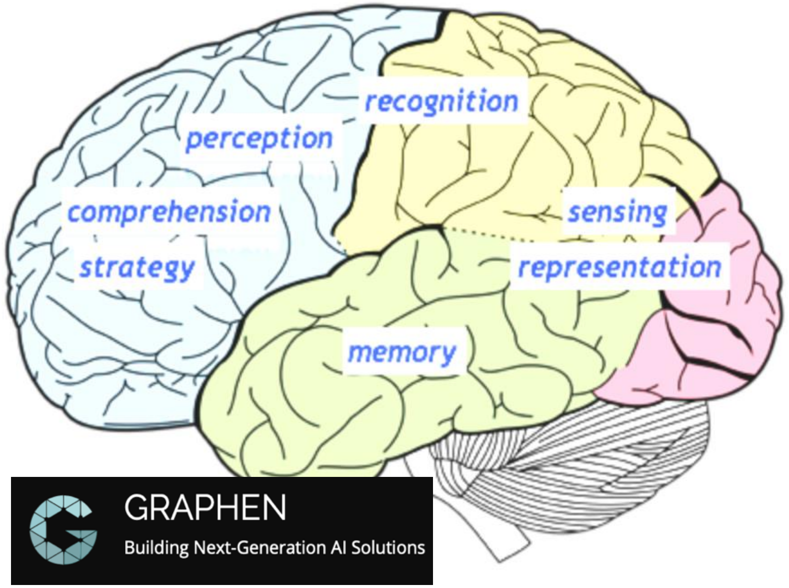
- Graphen Automobile – working with the North America market dominating leader to develop Car Doctor AI technology for the world

- Graphen AI was able to achieve almost 99% accuracy of car diagnosis and fix suggestions.

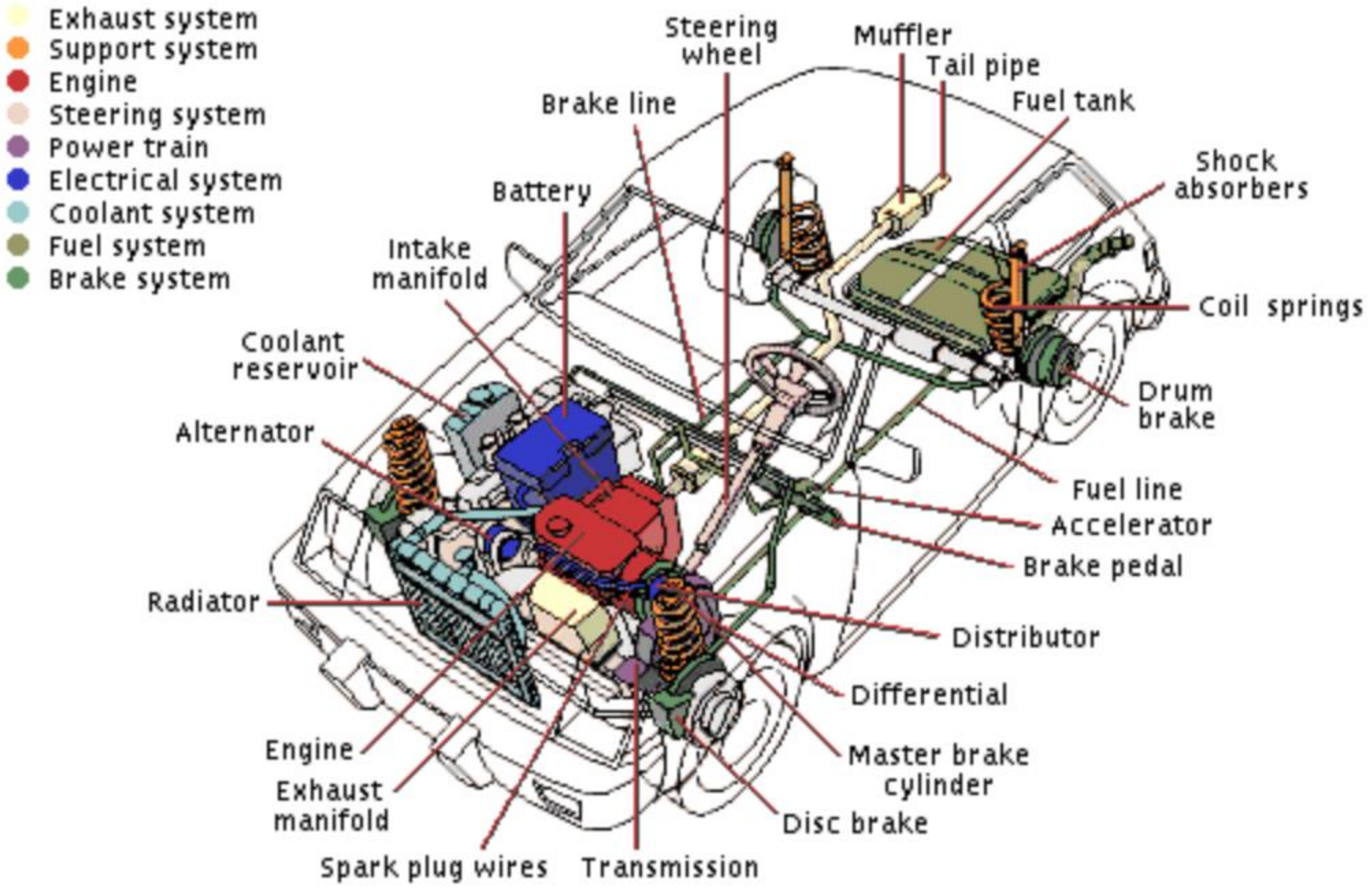
- Graphen AI was able to achieve 91% accuracy of car diagnosis and fix suggestions with just 1% of training data..

FixName1	Total present invalidation data	Total correctlypredicted by model	Accuracy Percentage
Inspect Cooling System and Repair As Necessary	77	77	100
Inspect Engine Oil Level and Fill or Replace as Necessary	285	285	100
Repair Engine Wiring Harness	59	59	100
Repair Fuel Injector Wiring	8	8	100
Repair Ignition Coil Wiring	184	184	100
Repair Mass Air Flow (MAF) Sensor Wiring	26	26	100
Repair Transmission Output Shaft Speed (OSS) Sensor Wiring	42	42	100
Repair Faulty Wiring in Engine Compartment	4	4	100
Replace Air Filter Element	299	299	100
Replace Camshaft Timing Gear	317	317	100
Replace Catalytic Converter(s) with new OE Catalytic Converter(s)	2960	2960	100
Replace Cylinder Head Temperature (CHT) Sensor	84	84	100
Replace Differential Pressure Feedback (DPFE) Sensor	4	4	100
Replace Electronic Oil Temperature Sensor (EOT)	243	243	100
Replace Engine Coolant Temperature Sensor (ECT)	363	363	100
Replace Evaporative Emissions (EVAP) Canister Vent Solenoid	160	160	100
Replace Evaporative Emissions (EVAP) Purge Solenoid	20	20	100
Replace Fuel Gauge Sending Unit	7	7	100
Replace Fuel Injector(s)	161	161	100
Replace Fuel Pump	190	190	100
Replace Fuel Pump Control Module	7	7	100
Replace Fuel Rail Pressure (FRP) Sensor	52	52	100
Replace Ignition Coil Boot(s) and Spark Plug(s)	15	15	100
Replace Intake Manifold Runner Control (IMRC) Actuator	224	224	100

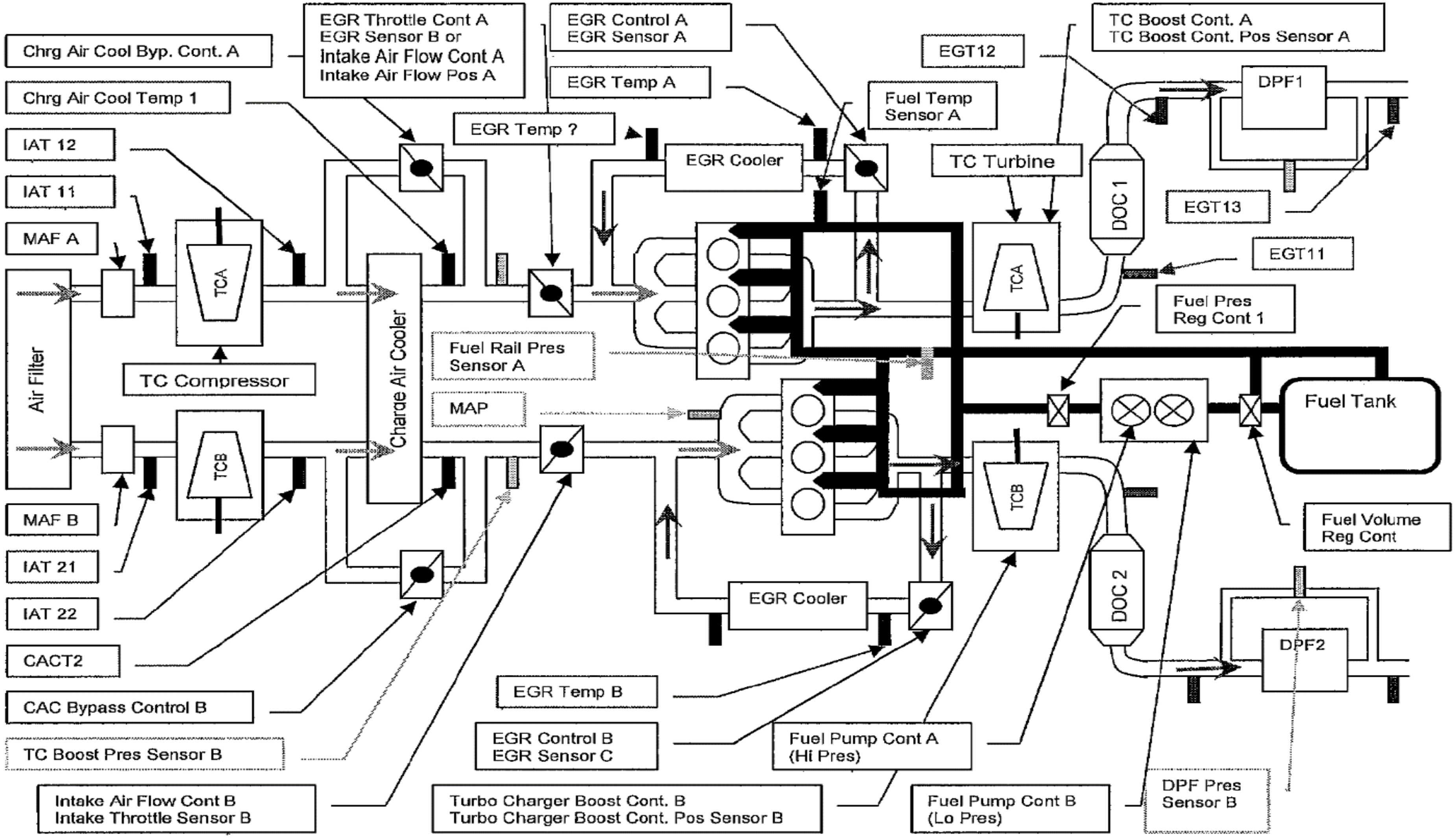
Graphen Automobile built Knowledge Graph of Cars

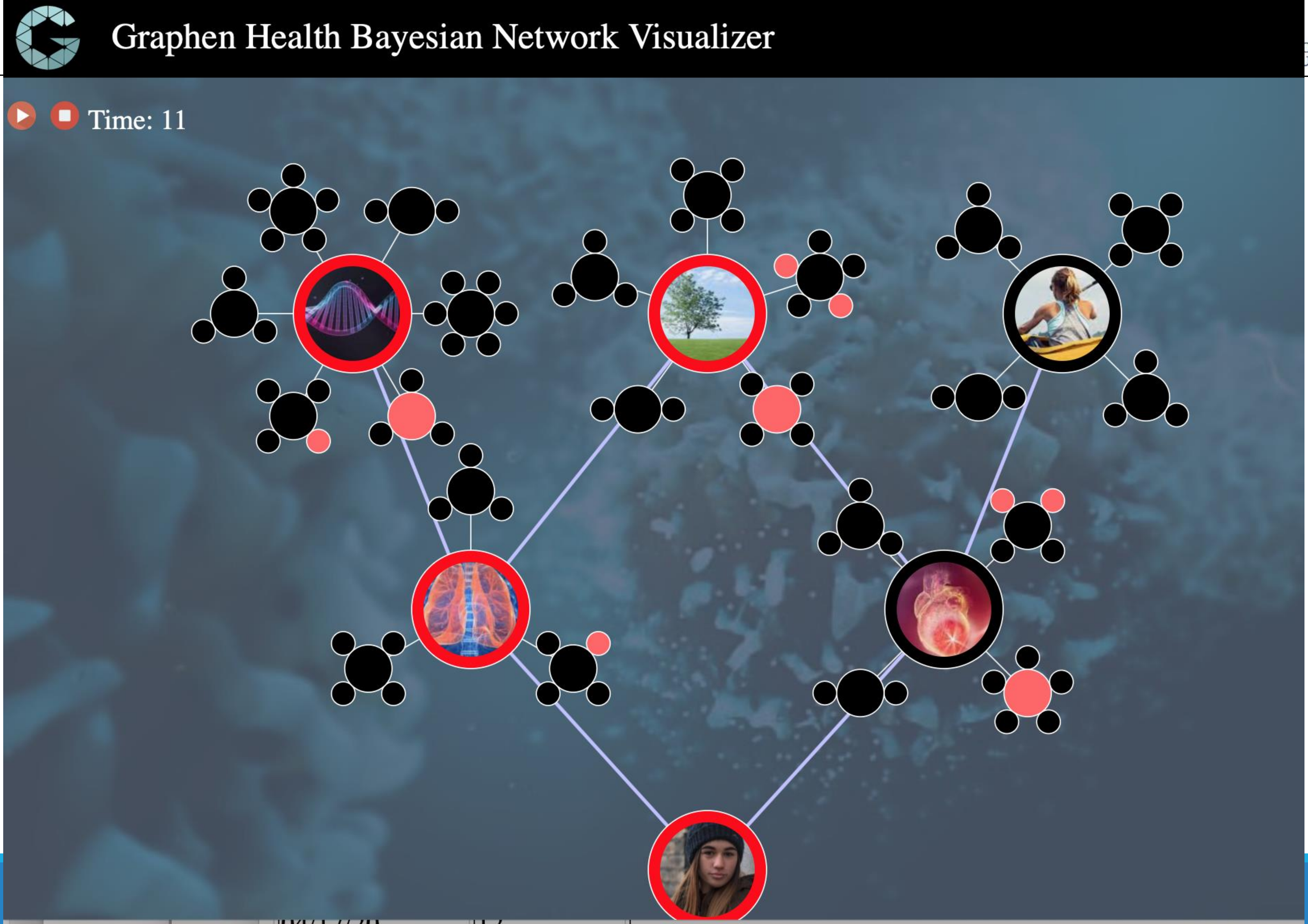


+



Graphen Automobile Car Sensor Knowledge Graphs

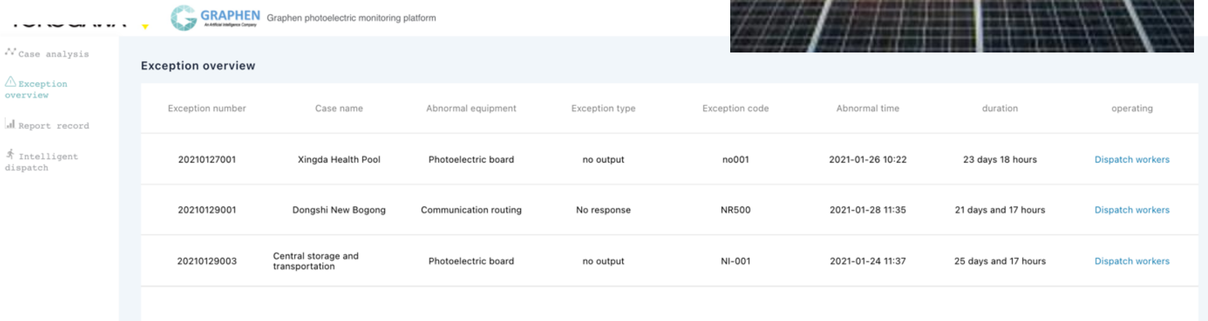



Graphen Health Bayesian Network Visualizer
Time: 11

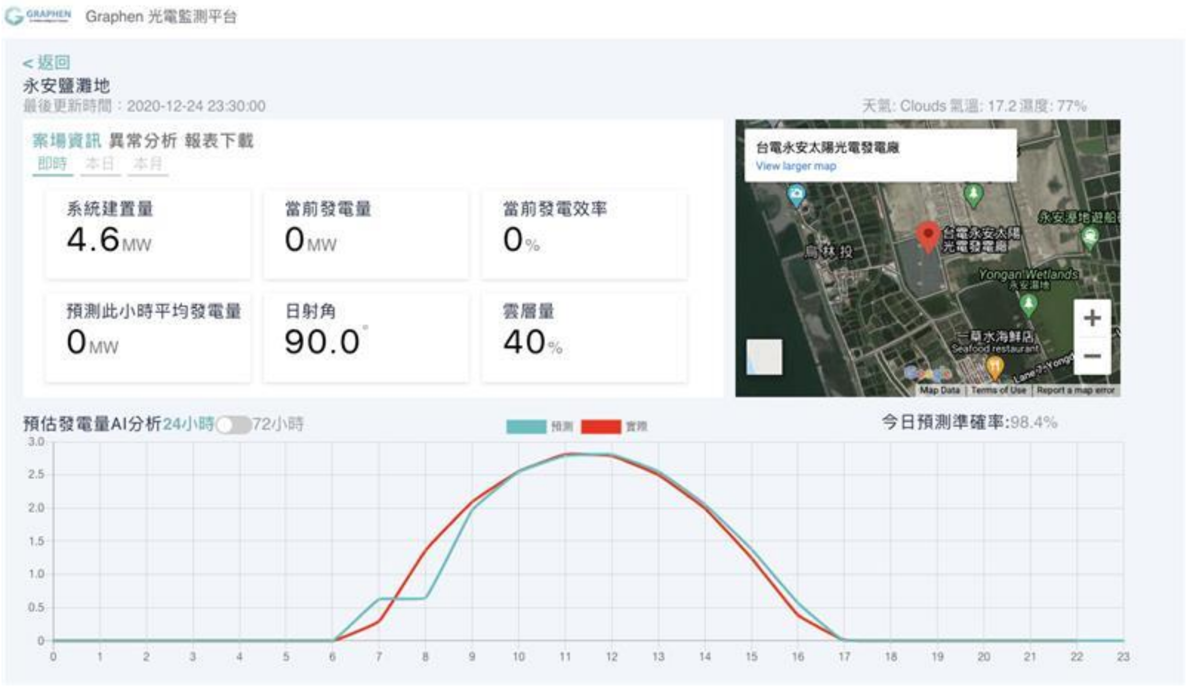
Ardi Applications – Graphen Energy

Graphen Energy – AI Reasoning & Strategy to realize Smart Grids

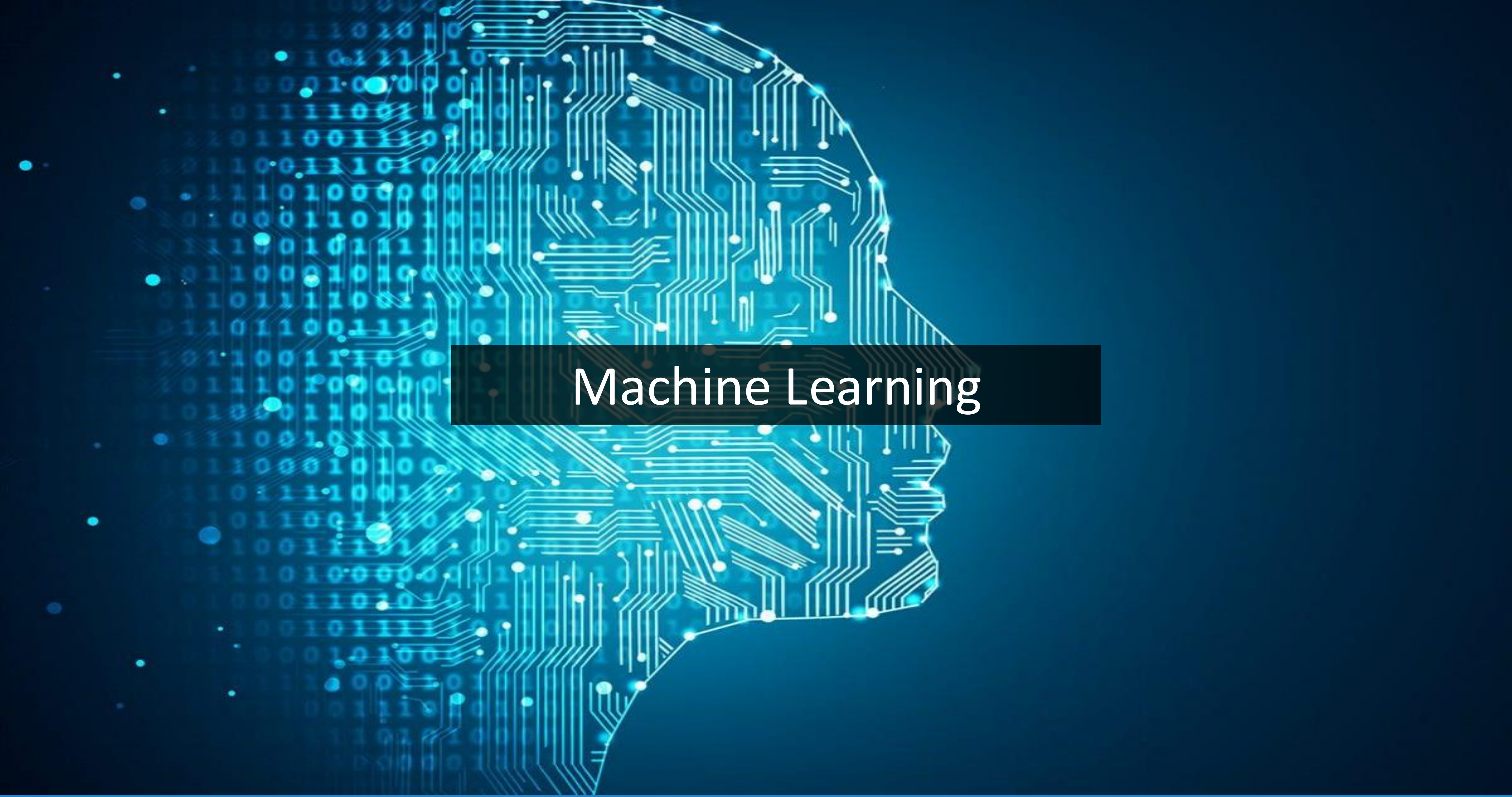
- **Renewable Energy Prediction**
- **Power System Anomaly Detection**
- Distributed Load Prediction
- Power Flow Analysis
- **Predictive Maintenance**



Exception number	Case name	Abnormal equipment	Exception type	Exception code	Abnormal time	duration	operating
20210127001	Xingda Health Pool	Photoelectric board	no output	no001	2021-01-26 10:22	23 days 18 hours	Dispatch workers
20210129001	Dongshi New Bogong	Communication routing	No response	NR500	2021-01-28 11:35	21 days and 17 hours	Dispatch workers
20210129003	Central storage and transportation	Photoelectric board	no output	NI-001	2021-01-24 11:37	25 days and 17 hours	Dispatch workers

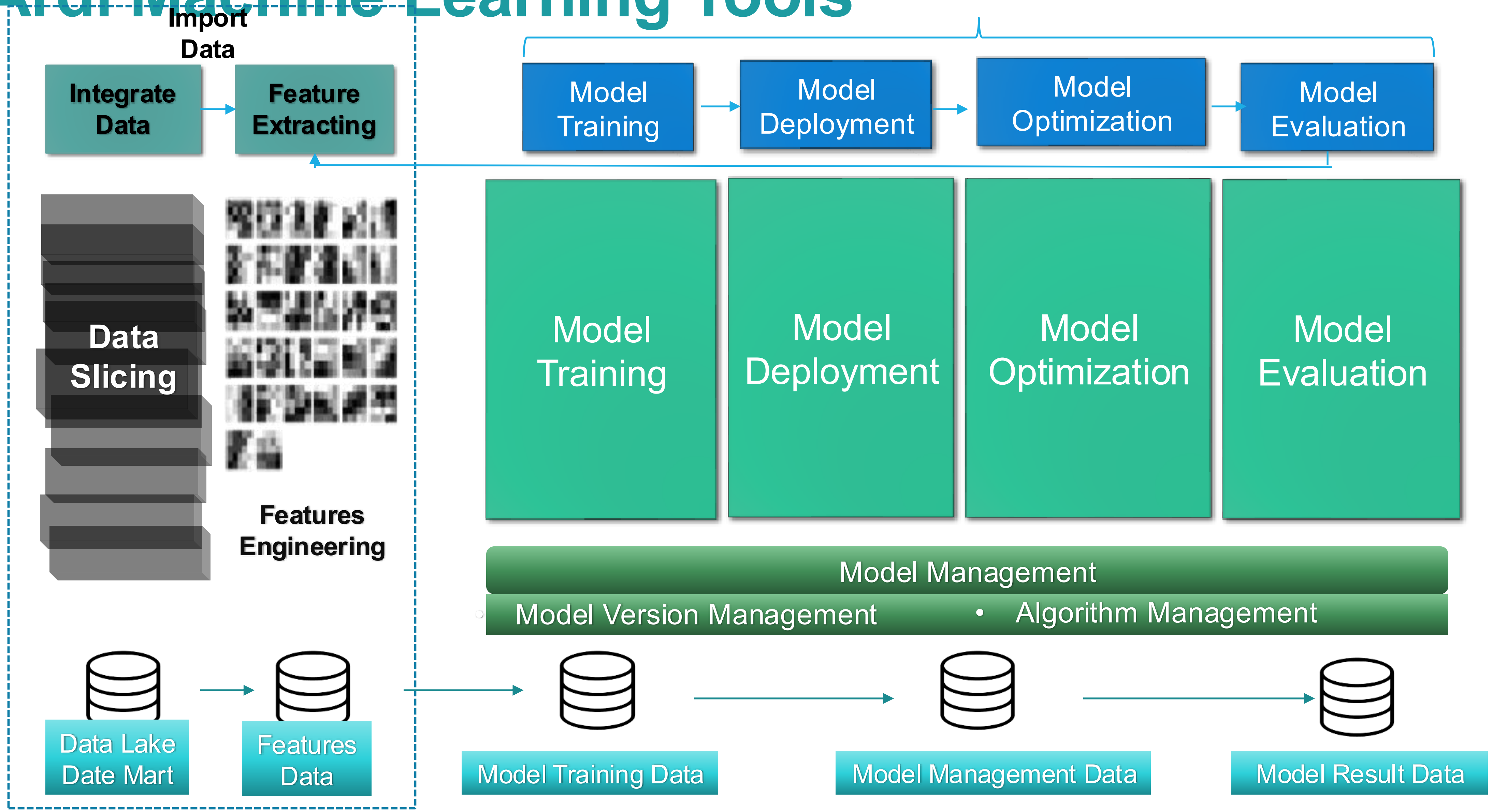


Graphen Energy's live system monitors all solar power stations in Taiwan and predicts power generations. Its accuracy is around 98.5%, far better than the customer's requirement of 90%.



Machine Learning

Ardi Machine Learning Tools



Ardi ML Algorithms Support

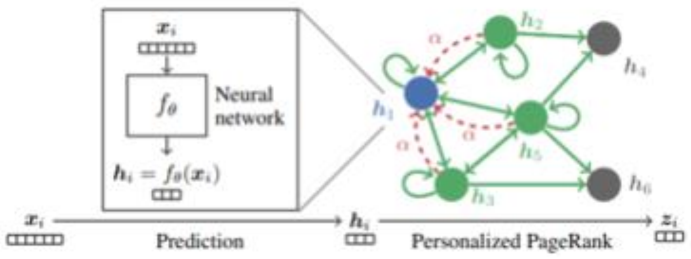
Classification
- Support Vector Machine
- XGBoost
- LightGBM
- Random Forest
- Decision Tree

Regression
- Ordinary Linear Regression
- Ridge Regression
- Logistic Regression

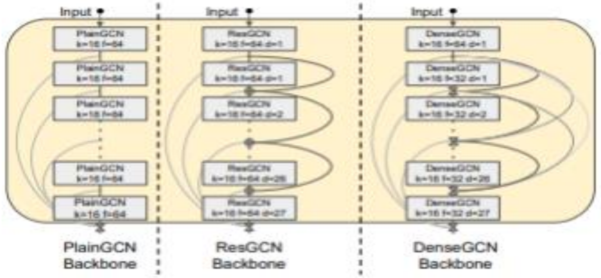
Clustering
- K-means
- Birch

Deep Learning
- Insert/delete layers
- Recurrent Neural Network (RNN)
- Deep neural network (DNN)
- Convolutional neural network (CNN)
- Graph neural network

Principal neighborhood aggregation



Approximate personalized propagation

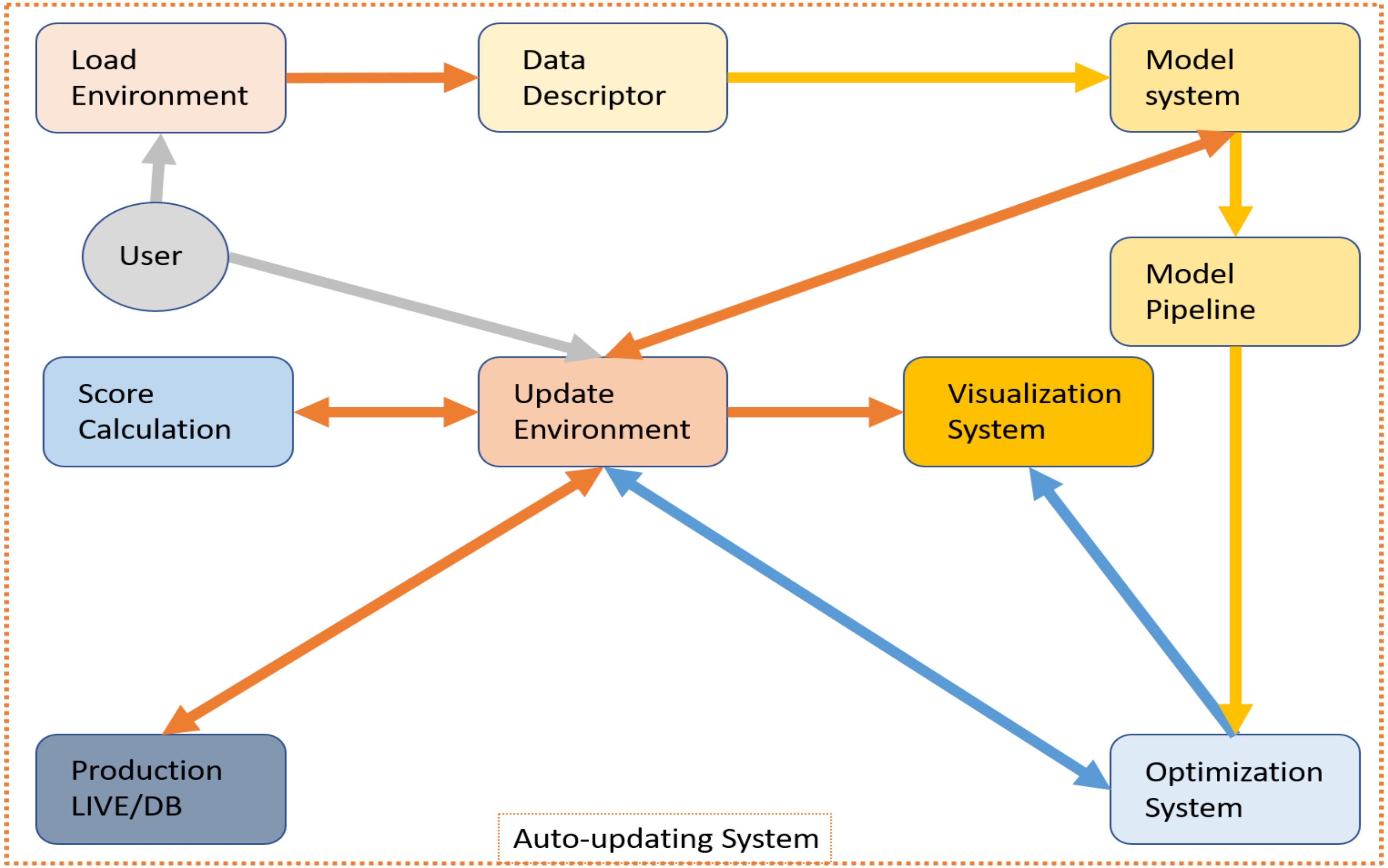


Deep GCN architecture

Ardi ML Modules

- **Model Training**– Provides convenient functions such a features importing and preprocessing to model developer， user can choose machine learning model and algorithm and tune parameters flexibly
- **Model Deployment**– Support user to set the frequency of model execution， model deployment and execution。
- **Model Optimization** – Support users to optimize model flexibly
- **Model Evaluation**– Support general evaluation criteria to regression and classification like accuracy and recall。
- **Model Management**– Integrated supports of importing various features data, choosing model type, saving model, setting access right and deployment. Support importation of models trained on outer platforms. Support automatically generate version of models.

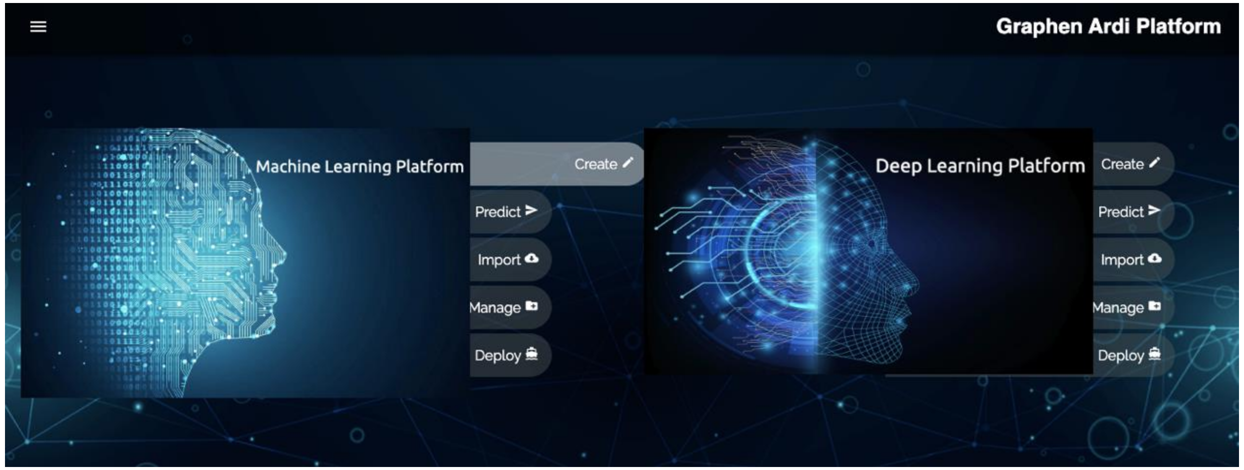
Ardi Automatic ML Optimization

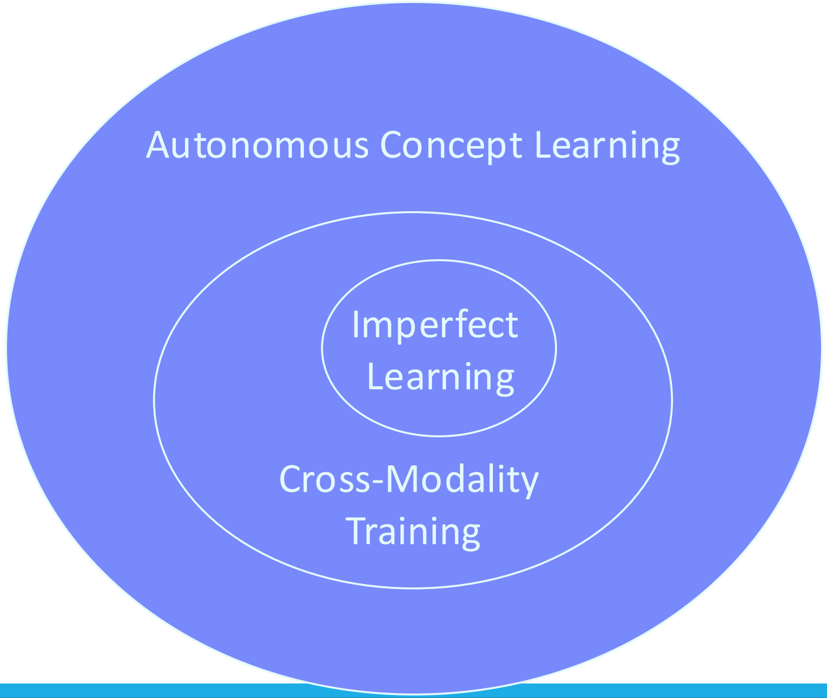


Autonomous Learning

Example: Autonomous Learning through Imperfect Training Labels

- Developed Machine Learning theories and algorithms for supervised concept learning from imperfect annotations -- imperfect learning
- Developed methodologies to obtain imperfect annotation – learning from cross-modality information or web links
- Developed algorithms and systems to generate concept models – novel generalized Multiple-Instance Learning algorithm with Uncertain Labeling Density





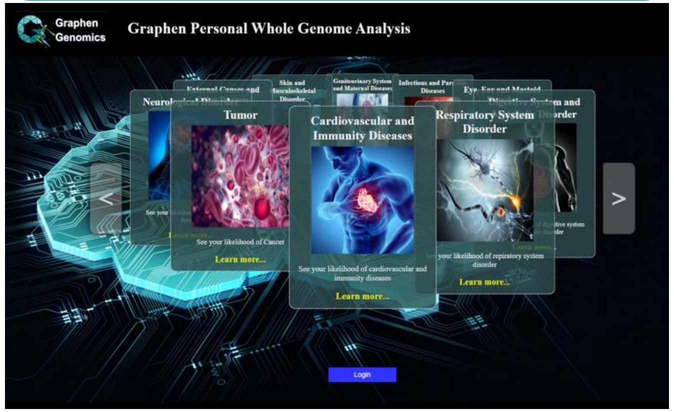
Ardi Applications – AI Medical

Columbia University

Graphen Medical – AI Meets the Central Dogma of Biology

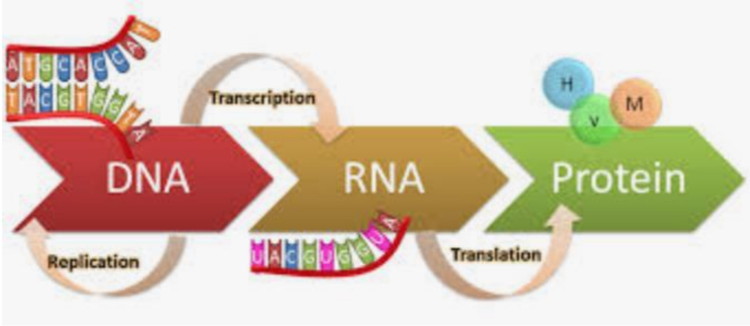
Personalized Whole Genome Disease Analysis

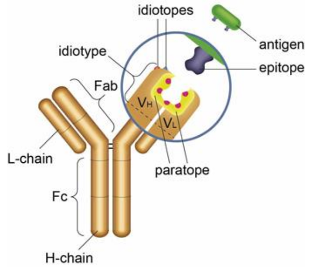
Large-Scale AI Medical Article Understanding

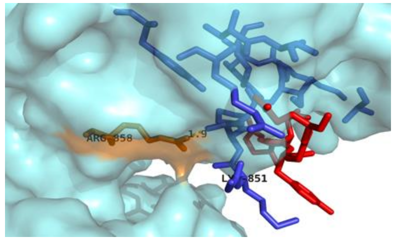


- Utilizing AI technologies to read tens of thousands of medical articles;
- Combining with Whole Genome Sequencing of 3.2B pairs of human genome;
- Predicting risks of ~400 diseases

AI Tools for Drug Development



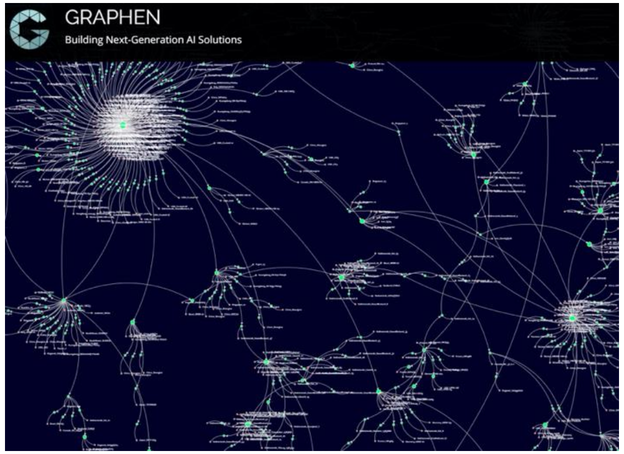




- Using AI to build Protein Structure and Function prediction models, and predict Drug Target Affinity, ADME, and Antibody/Antigen selection models

Virus Mutation Surveillance

Personalized Precision Medicine



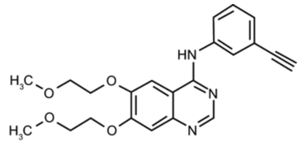
- Strain surveillance and mutation function prediction to the detail of countries, states, and cities.
- Disease progress prediction and personalized therapy solution suggestions.

Graphen AI Tools for Drug Development

Personal Druggable Candidates Finding 1

Well-Known Candidates 1

Candidates Preprocessing 2



Precision Drug Development 4

Small molecular/ Biosimilar Drug Development 3

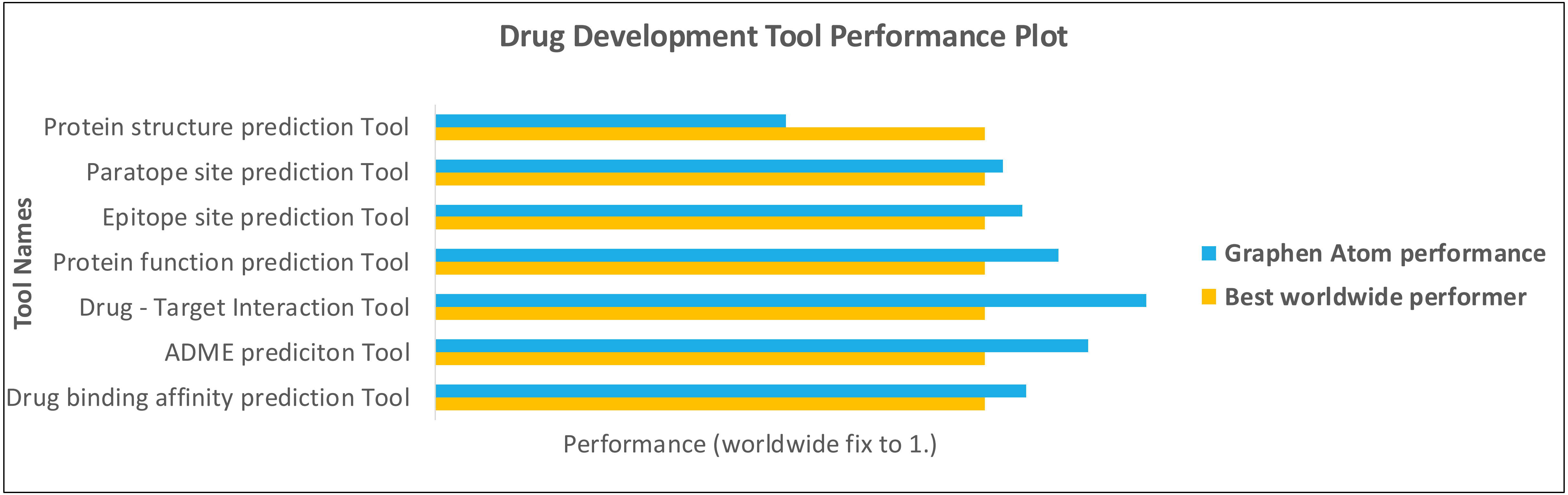
Precision Medicine

Common Drug Development

All but one Graphen Atom Drug Tools outperformed the bests in the world

Atom Protein Drug / Small Molecular Drug de novo Development Tools

- Computing requirements > 1 x Nvidia V-100 GPU (32 GB)
- Only comparing our tools when there are other tools to compare in the literature
- Graphen Atom (AI tools for Medicine) outperforms known best worldwide performers in all tools except the protein structure prediction Tool (by Google DeepMind).



Strategy

What-If Assessment

○ Holistic Approaches

- To analyze the customer's behavior vs. control groups:
  - Evaluate customer risk via analysis of customer behavior and relationship changes
  - Discrepancies with self or peer group behavior within the same industry
  - Detect anomalous transaction behavior, including frequency and suspicious counterparties

Machine Learning Solutions

- Time Series Analyses
- Graph Analysis
- Supervised Machine Learning:
  - Regression, Clustering, etc.
- Unsupervised Machine Learning:
  - Clustering, Local Outlier Factorization, etc.

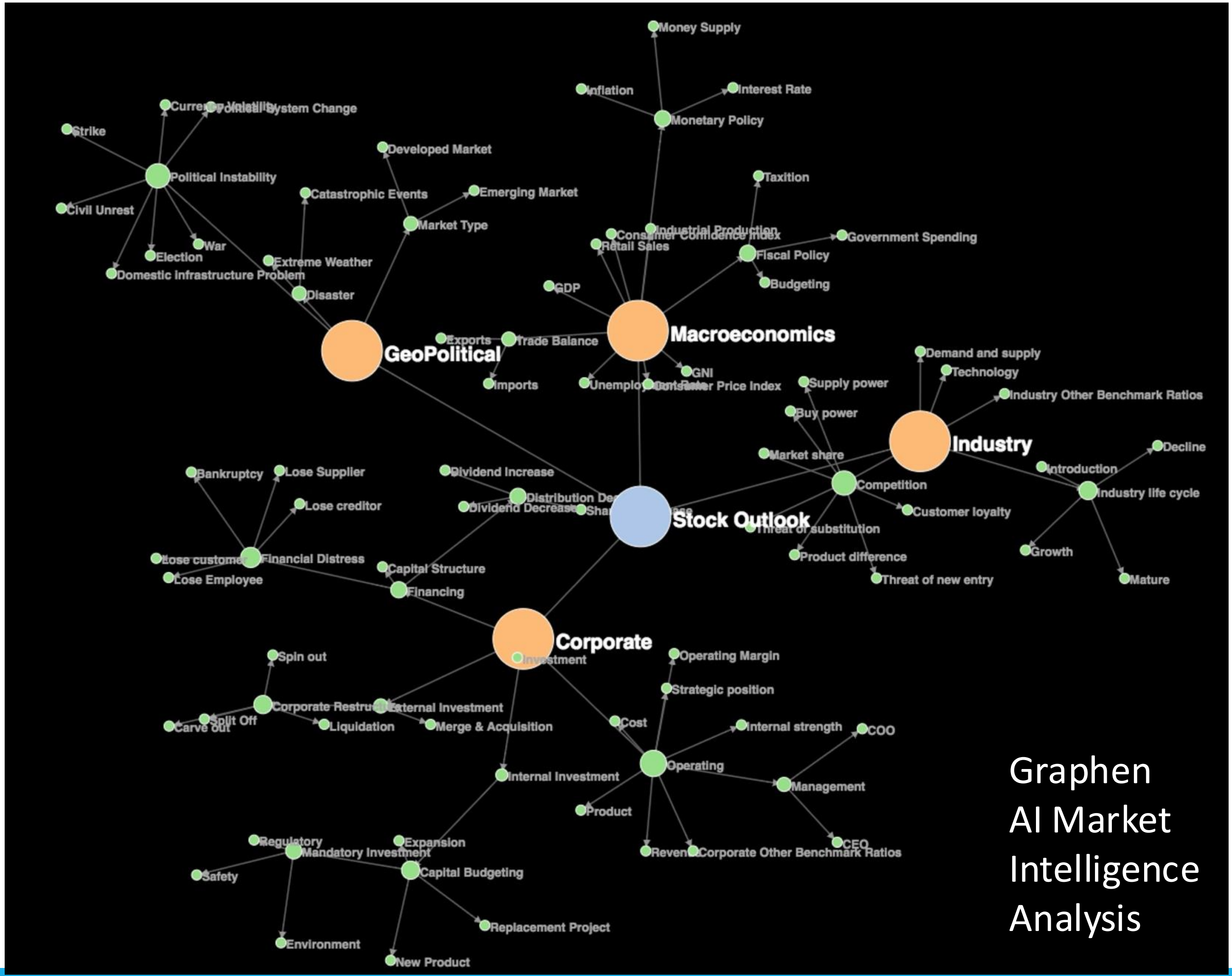
Underwriting criteria:

Pricing	Per CRE pric
Min. DSCR	1.25X 1.x ba appraiser's p
Stress test: current + 2%, 25-yr	1.0X Cash flow, i.e increasing op
Operating expenses	Based on the historical per expenses
Tenant estoppels	Per loan polic
Auto payment debit	Yes – per CR

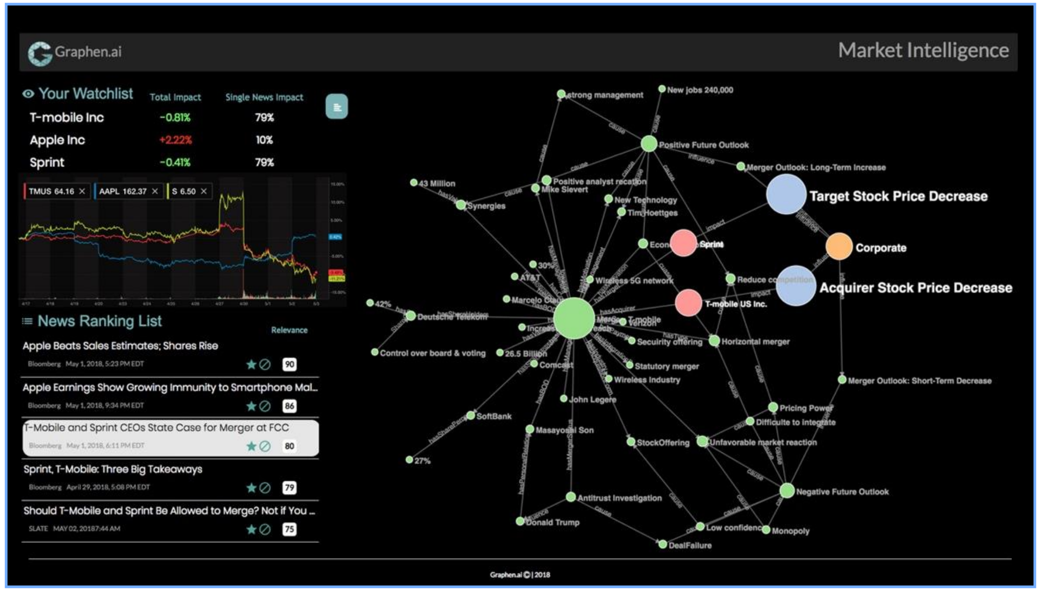
AI-powered risk scoring with continuous application monitoring

Scoring on various scenarios of stress tests and what-if conditions, assign predictions on incomplete fields, scores update as new info discovered or provided.

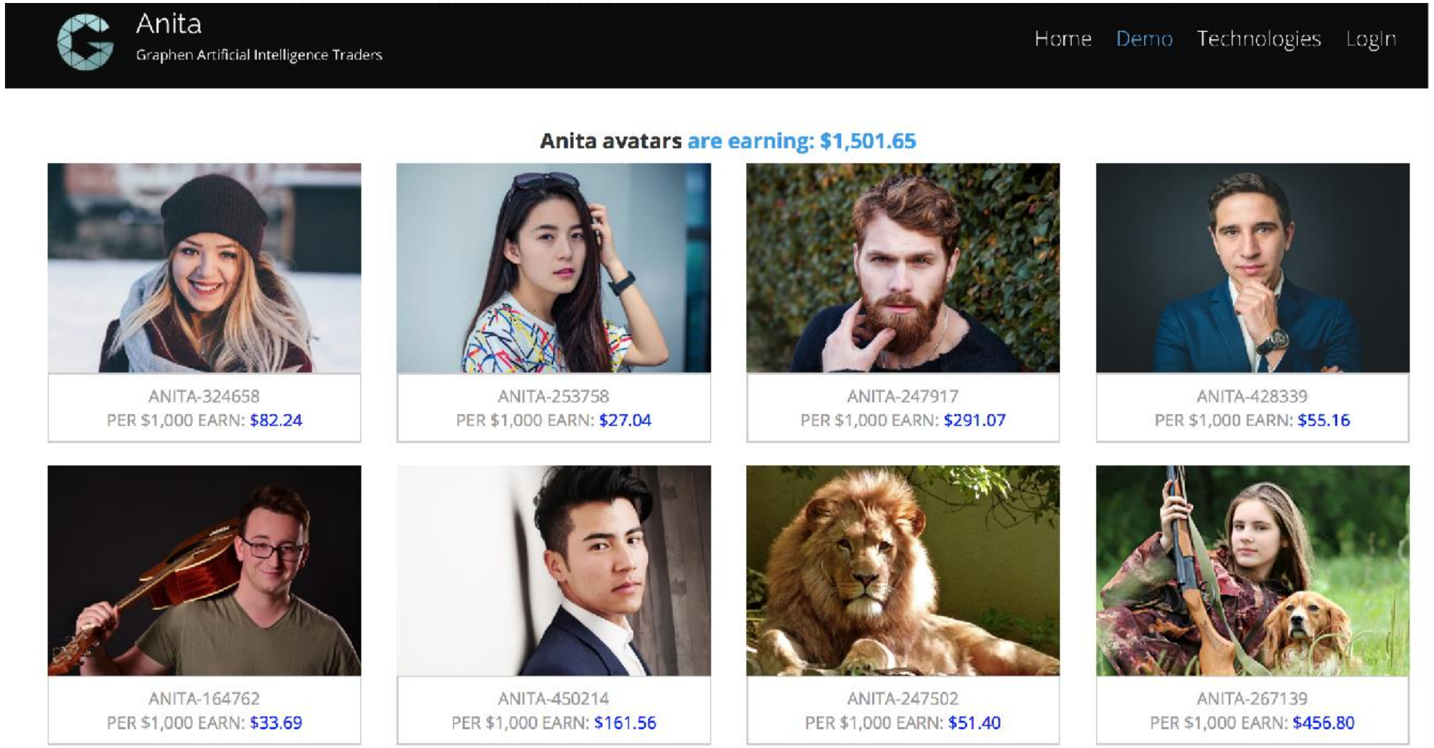
Example: Market Intelligence



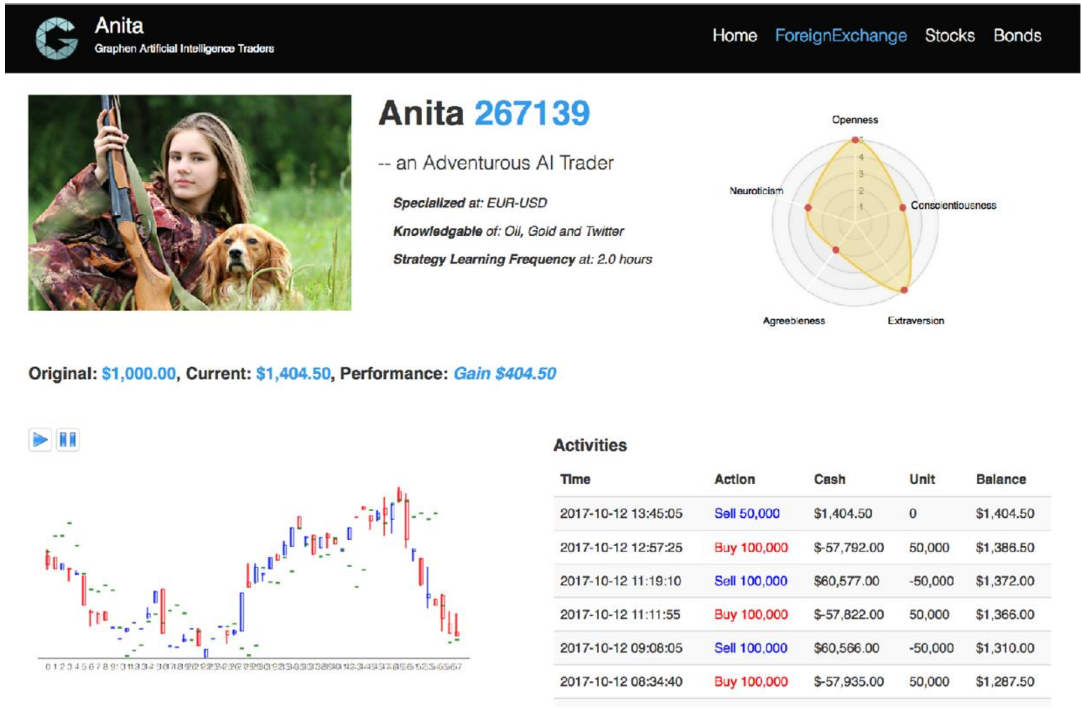
- Construct Factors that impact company's performance
- Construct the Influence Knowledge Graphs that interconnect between companies
- Simulate What-If Scenarios

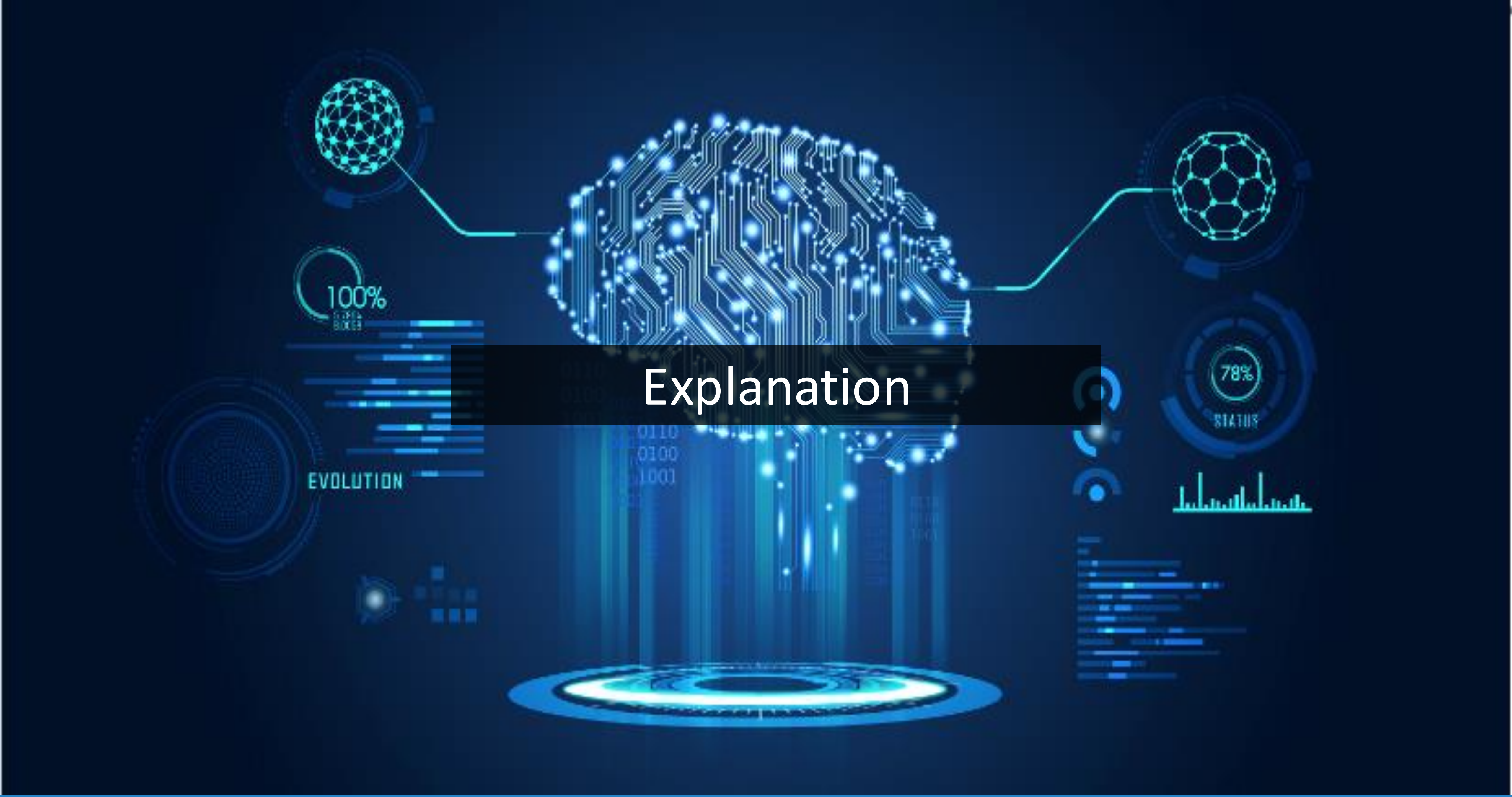


Example: AI Trader



- Avatars with different trading strategy
- Simulate personalities





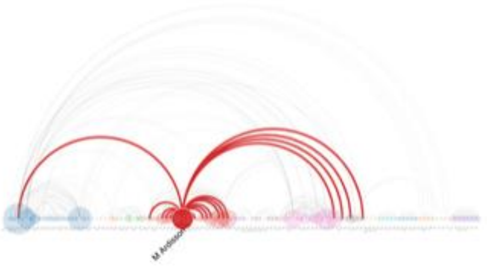
Explanation

Graphen AI Explainer

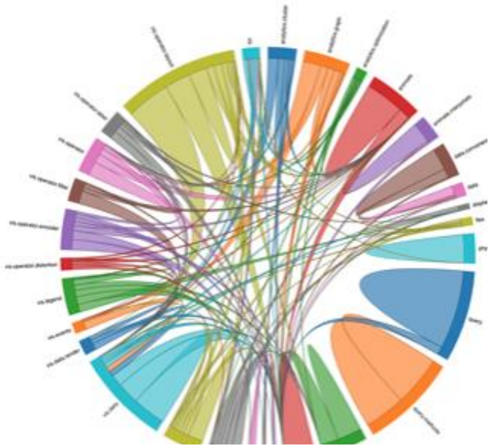
- Local interpretable model-agnostic explanations
- Approximates the model to closest linear model at a local level
- Takes a point and generates several point in vicinity
- Generates a linear model based on result of above points
- Calculates the weightage of each input feature

Graphen Visualization Tools

Arc Diagram Hover:

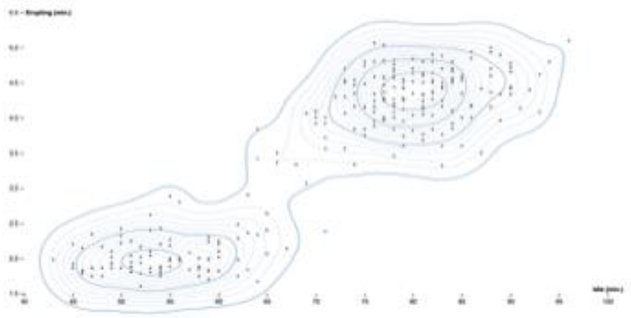


Chord Dependency Static:



Tree Map Static:

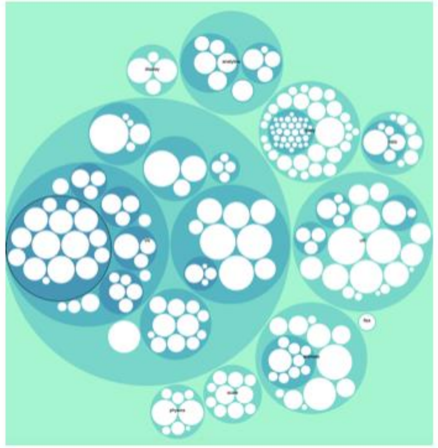
Density Contours Static:



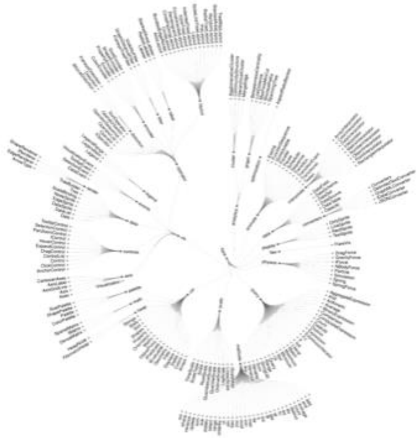
- 30+ different types of graphics

Sunburst Zoomable:

Circle Packing Zoomable:

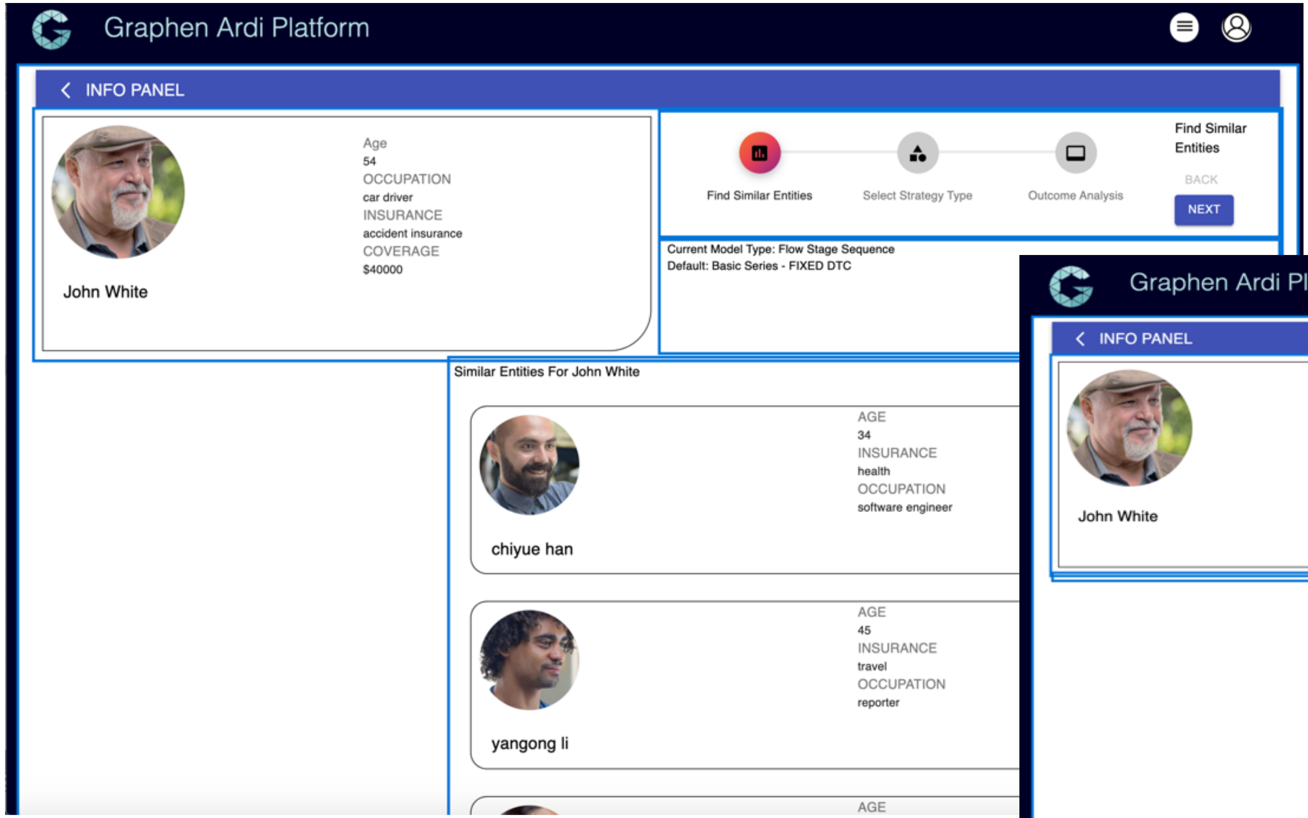


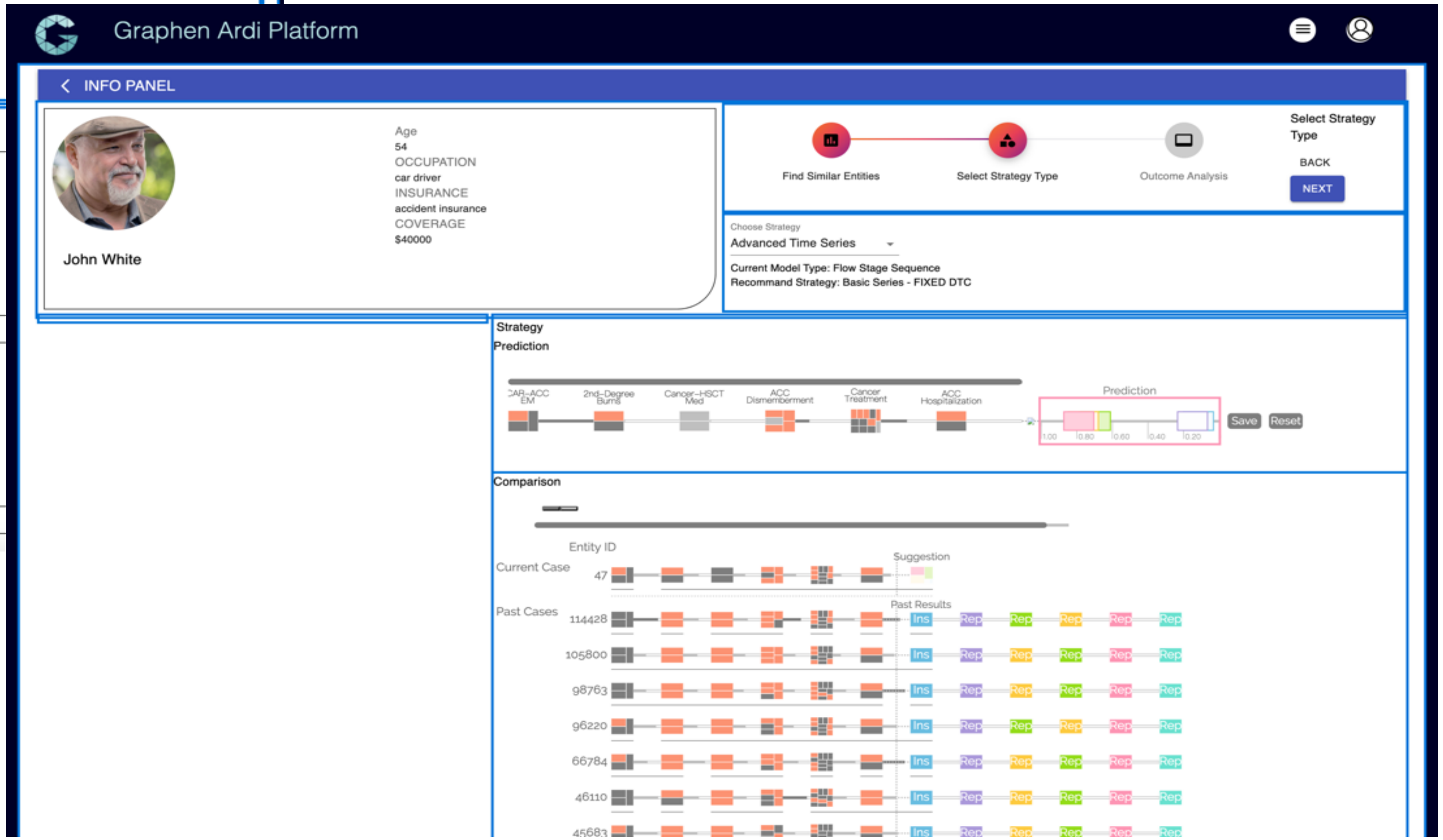
Radial Tidy Tree Static:



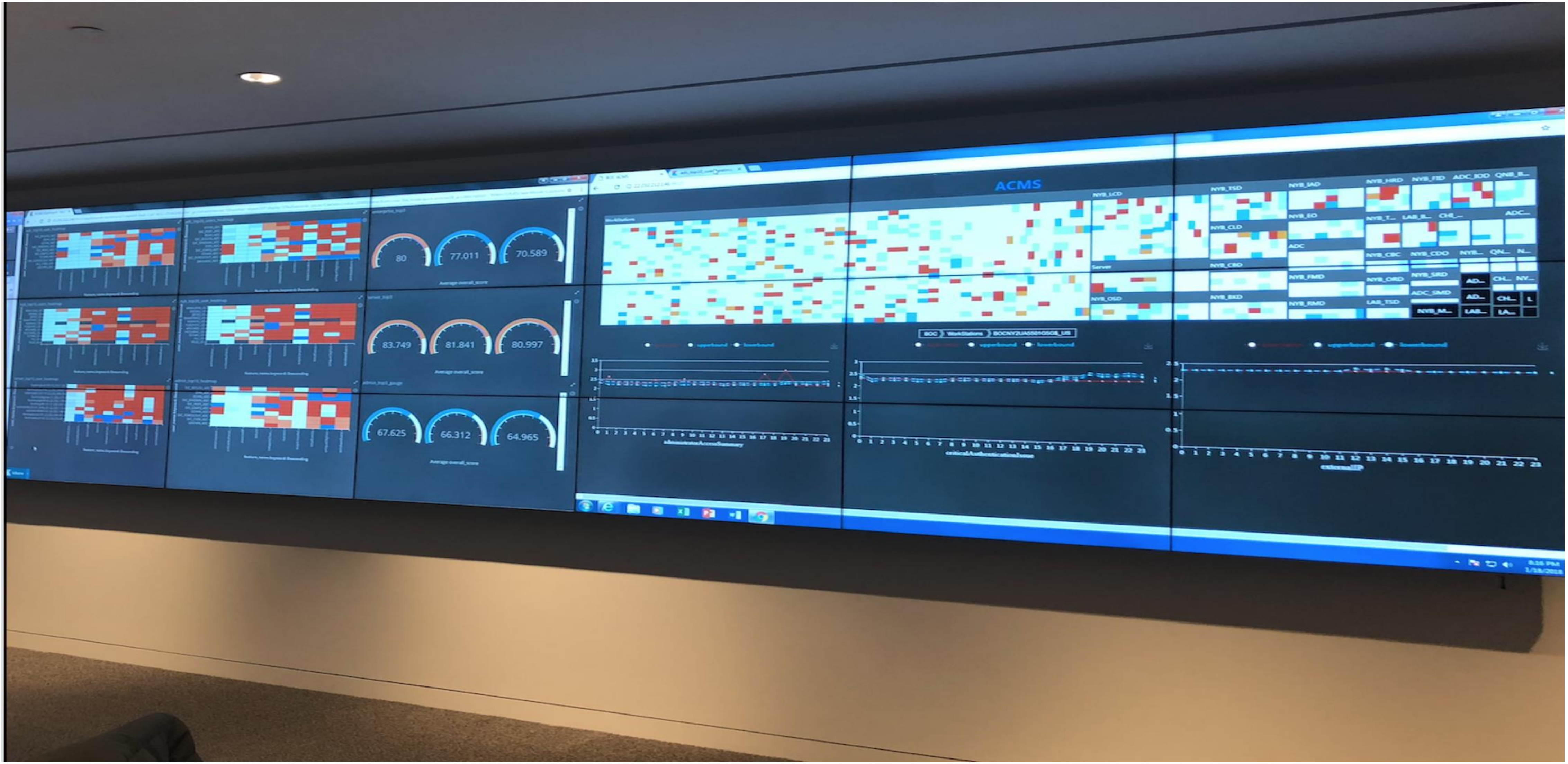
Worldmap with Tracing Bar Dynamic:

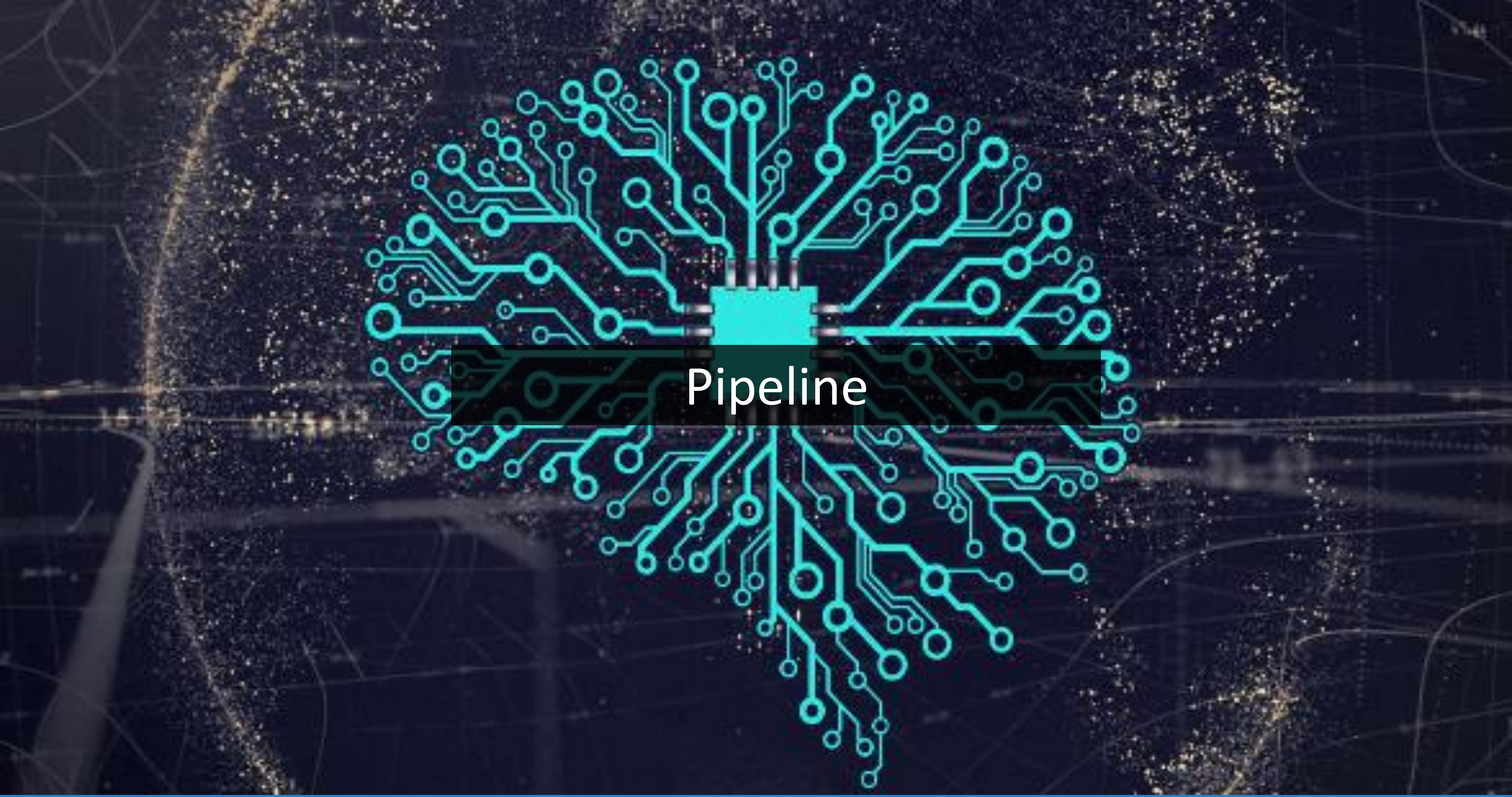
Ardi Explanation – Health Monitoring





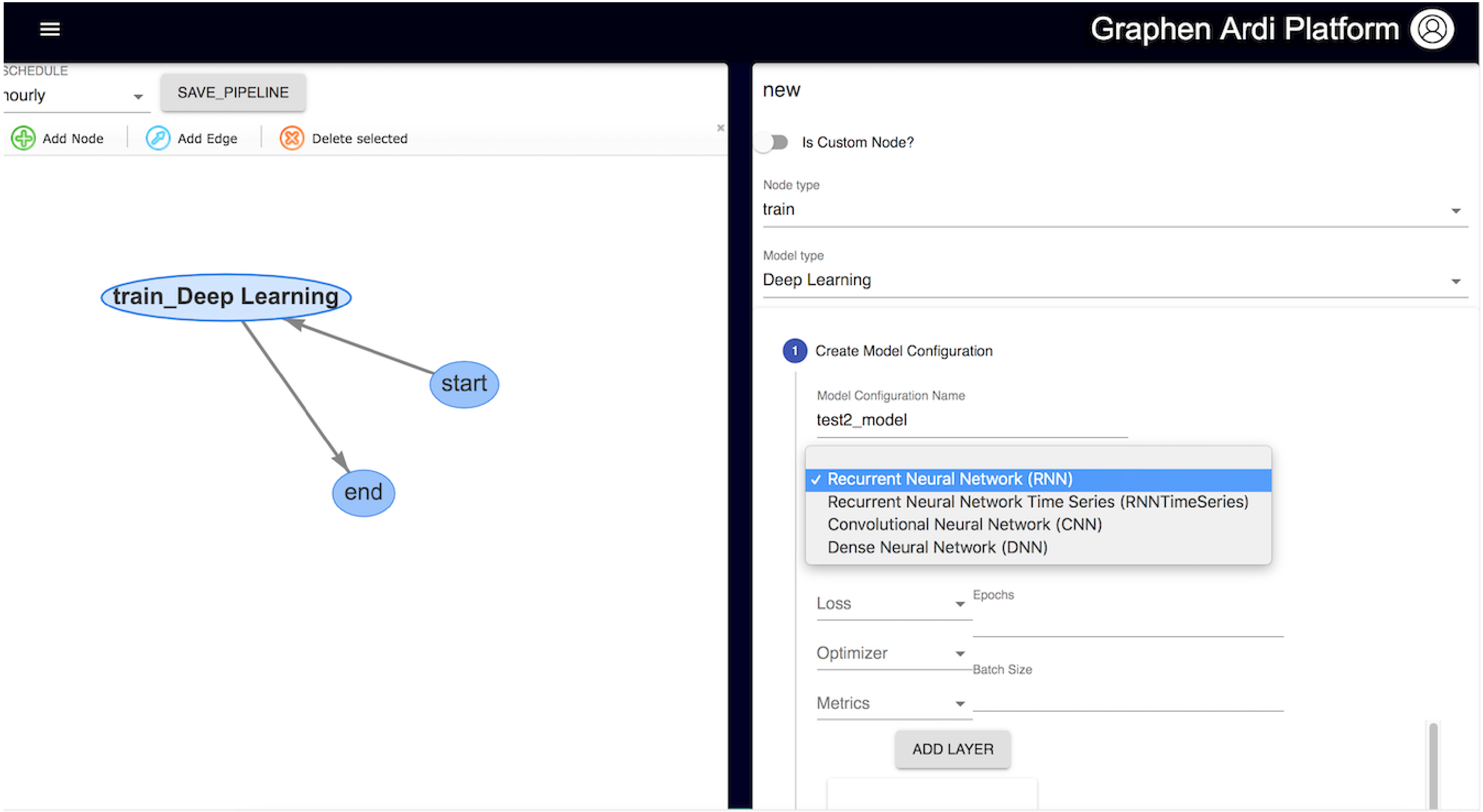
Example: Core-Banking Monitoring Center





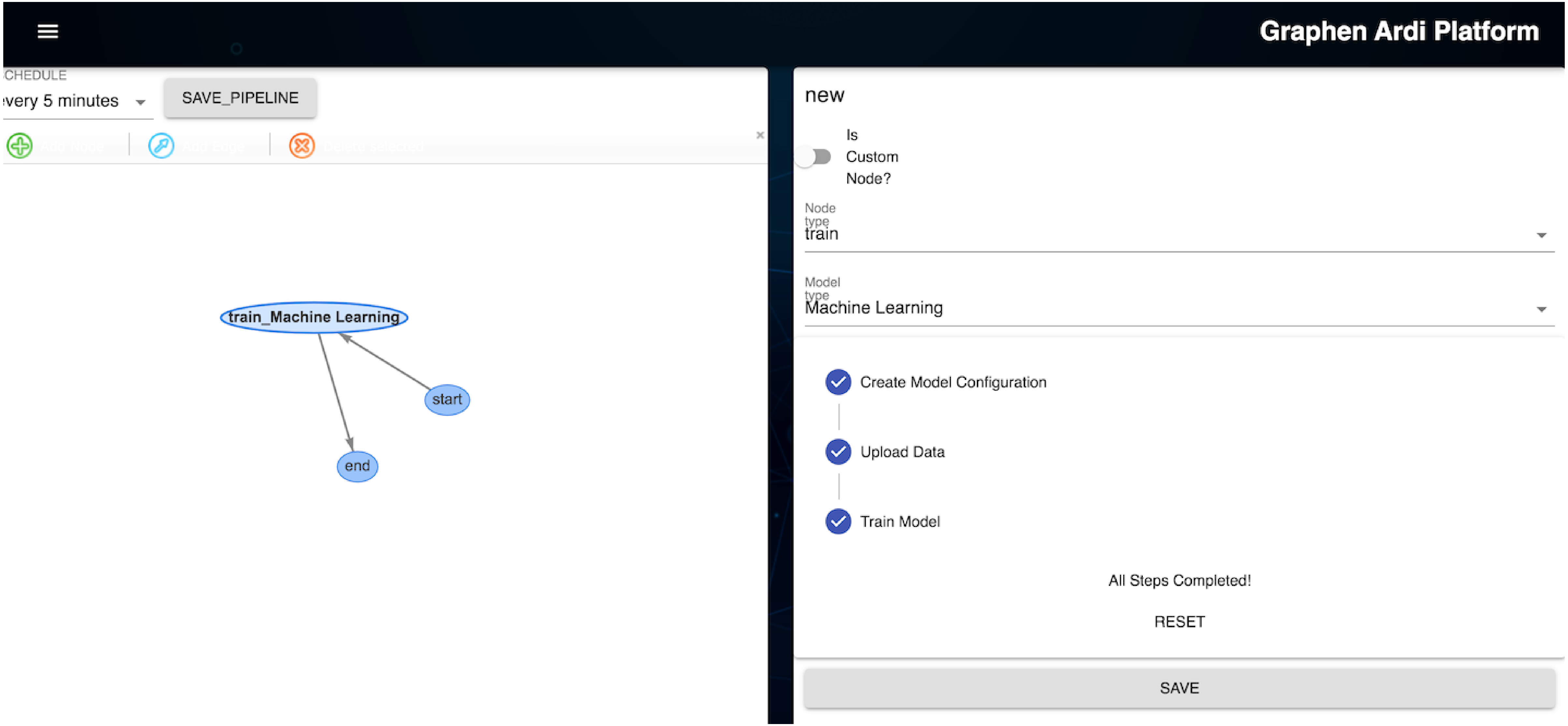
Pipeline

Ardi Pipeline Tools

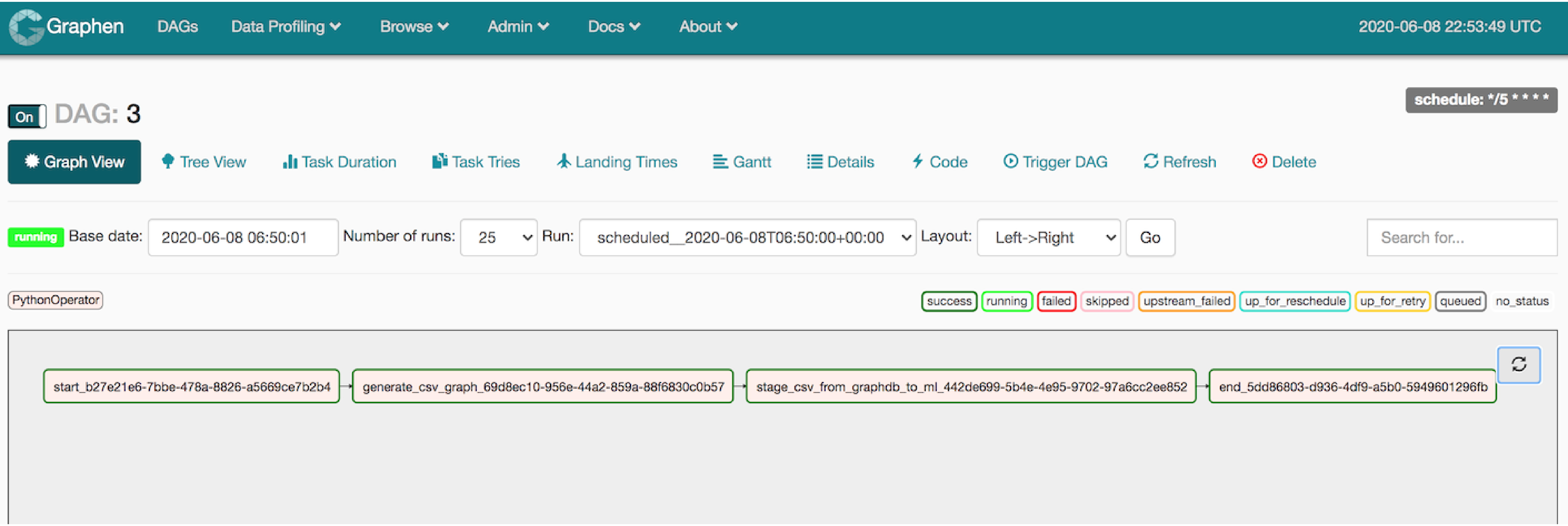


- Scheduling tasks
- Cascading task steps
- Publish for production applications
- Monitoring status of steps

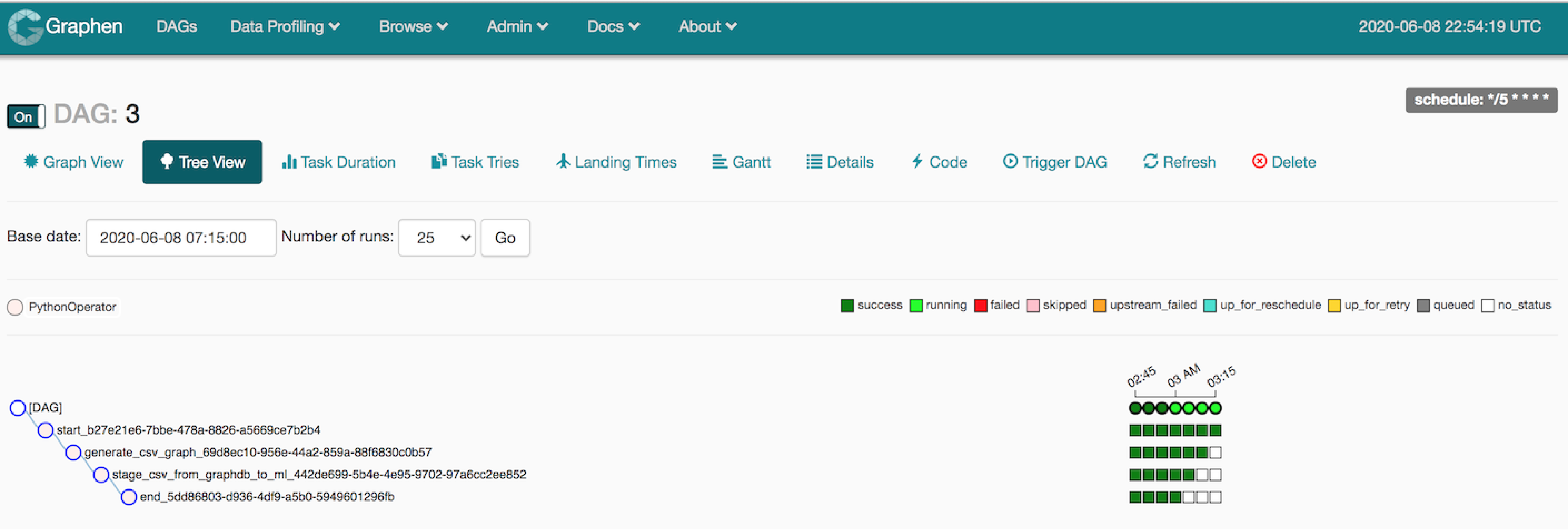
Ardi Pipeline Tools



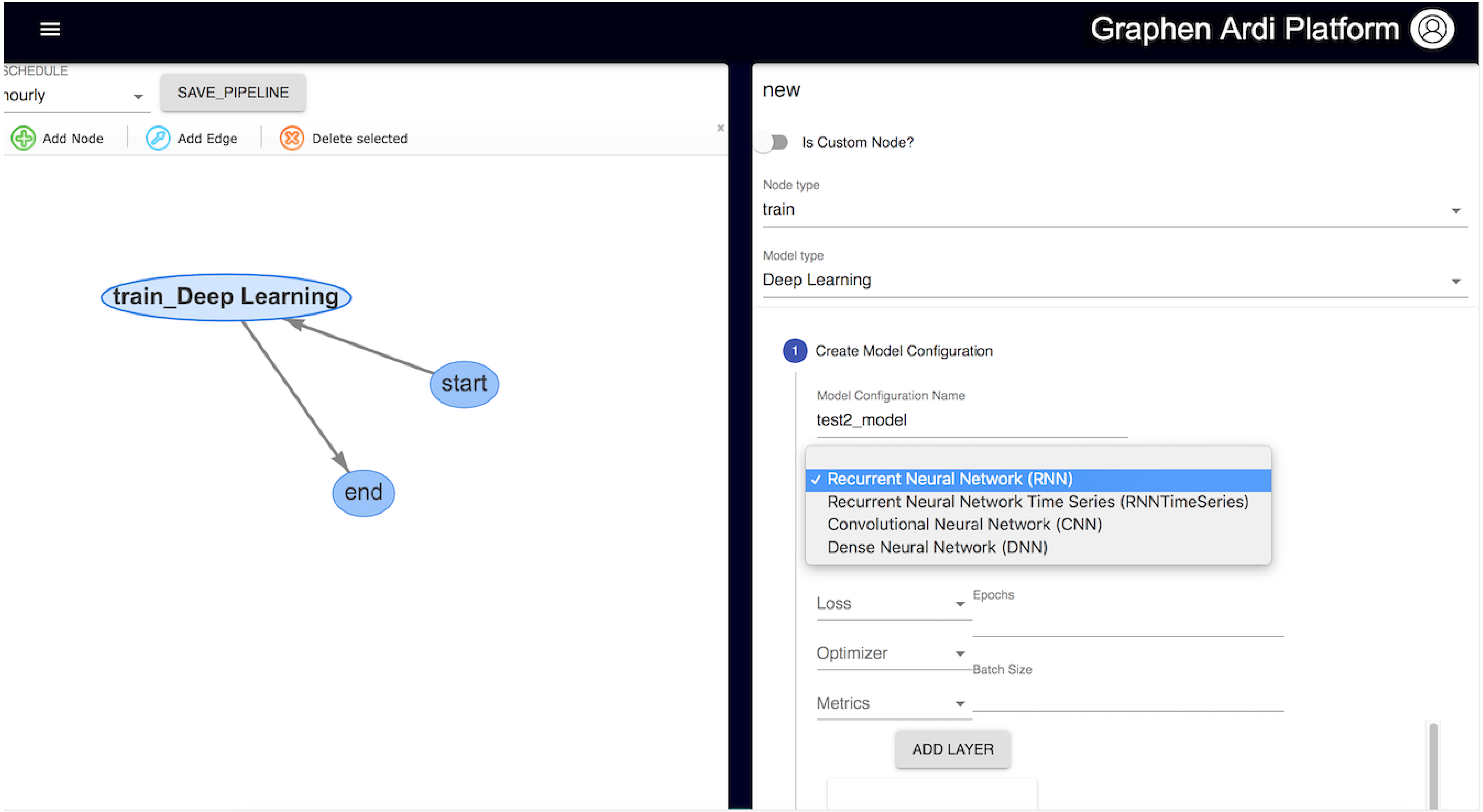
Ardi Pipeline Tools



Ardi Pipeline Tools



Ardi Pipeline Tools

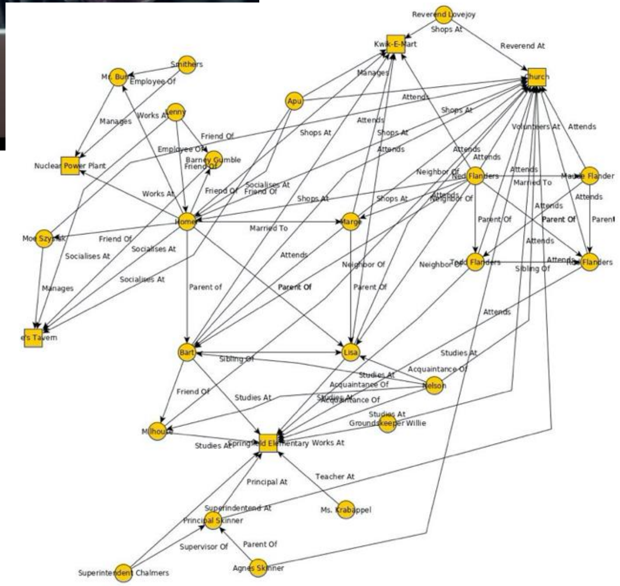


- Scheduling tasks
- Cascading task steps
- Publish for production applications
- Monitoring status of steps

Sense

Ardi Sense

Deep Video Understanding + Natural Language Understanding



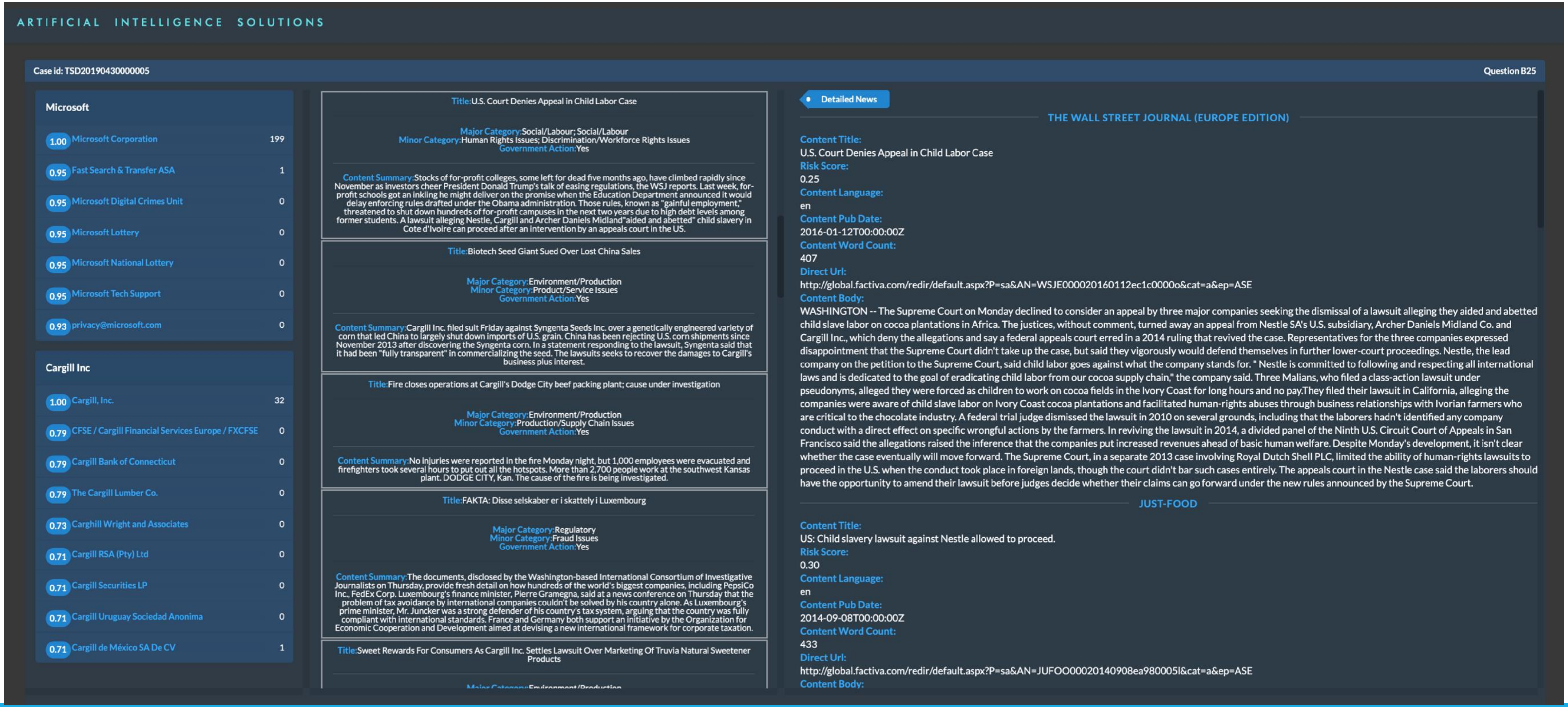
Graphen’s Ardi Sense achieves Deep Video Understanding in the ACM Multimedia 2020 Grand Challenge (2nd place):

- Visual Recognition
- Speech Recognition
- Knowledge Graph
- Face Recognition
- Emotion Recognition
- Speaker Identification
- Relationship Inference
- Event and Action Understanding

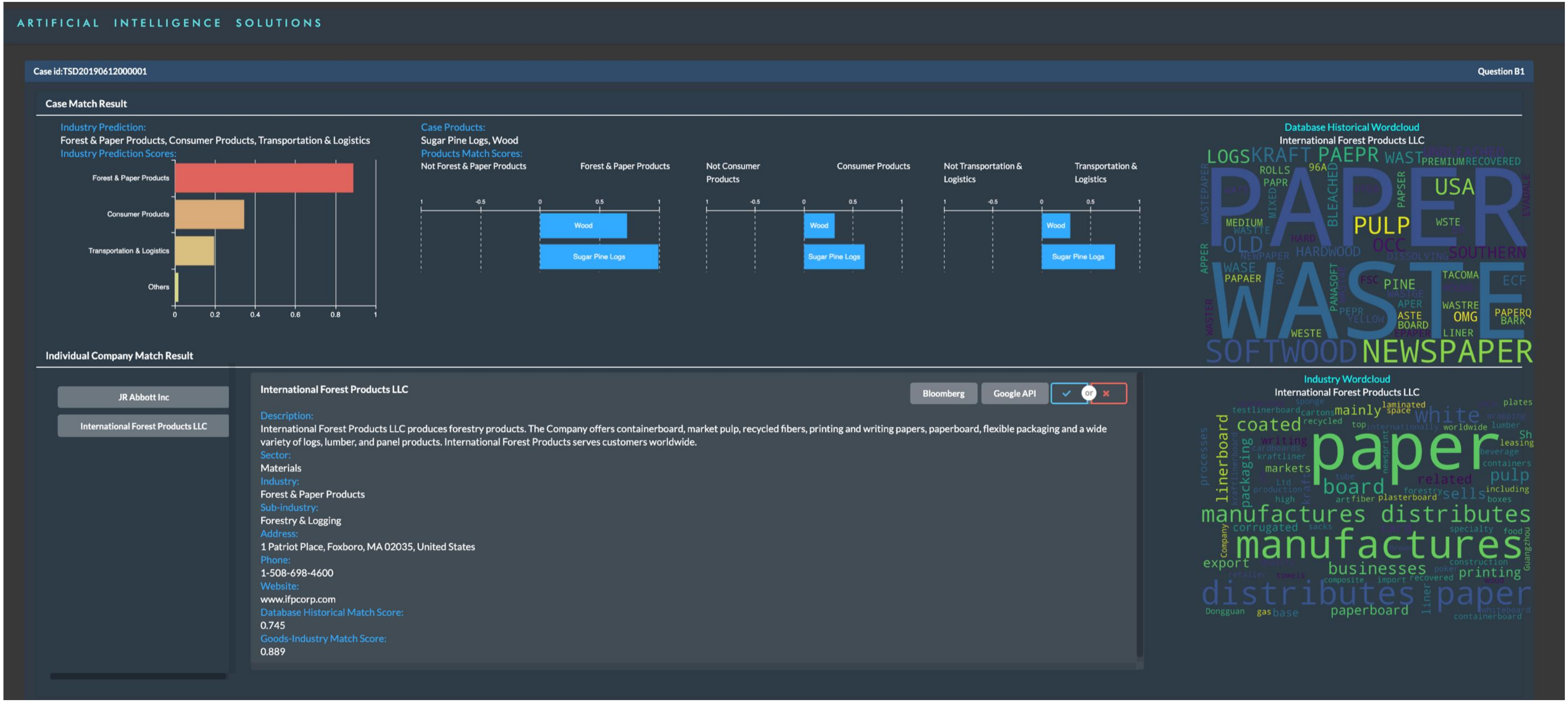
Ardi Sense’s Natural Language Understanding:

- Understand unstructured text within context
- Reading Medical Articles -> summarization & Q&A
- Reading Financial Information and Market Info

Example: Negative News



Example: Company Due Diligence



Graphen Personal Whole Genome Analysis

Graphen Whole Genome Analysis

Do you want to know yourself? What does your blueprint say about you? Graphen Personal Whole Genome Analytics System analyzes your entire 6.4B genome. It provides your risk likehood of 350+ diseases in 10 major categories:

- **Tumor**
- **Cardiovascular and Immunity Diseases**
- **Respiratory System Disorder**
- **Digestive System and Metabolic Disorder**
- **Eye, Ear and Mastoid Disorder**
- **Infectious and Parasitic Diseases**
- **Genitourinary System and Maternal Diseases**
- **Skin and Musculoskeletal Disorder**
- **Neurological Disorder**
- **External Causes and Other Diseases**

Ardi Functions Recap

- Graph Database
- Relational Database

Database

Analytics

- Graph Analytics
- Feature Engineering

Learning

- Machine Learning
- Deep Learning
- Autonomous Model Optimization

Sense

- Natural Language Processing
- Deep Video Understanding

Pipeline

- Production Workflow

Reasoning

- Causality Modeling
- Behavior Prediction

Explanation

- Visualization
- ML Explanations

Strategy

- Action Strategy Simulation

Ardi AI Platform

Summary

- A suite of AI powered offerings from foundational platform to industry applications

AI Foundation | Full-Brain Platform

AI Finance | Risk, Fraud, ESG & Intelligence

AI Medical | Knowledge, Drugs & Precision

AI Automobile | Car Doctor

AI Energy | Clean Energy & Smart Grid

AI Ethics

Google DeepMind

2024-04-19

The Ethics of Advanced AI Assistants

Iason Gabriel[*,1], Arianna Manzini[*,1], Geoff Keeling[*,2], Lisa Anne Hendricks[1], Verena Rieser[1], Hasan Iqbal[1], Nenad Tomašev[1], Ira Ktena[1], Zachary Kenton[1], Mikel Rodriguez[1], Seliem El-Sayed[1], Sasha Brown[1], Canfer Akbulut[1], Andrew Trask[1], Edward Hughes[1], A. Stevie Bergman[1], Renee Shelby[2], Nahema Marchal[1], Conor Griffin[1], Juan Mateos-Garcia[1], Laura Weidinger[1], Winnie Street[2], Benjamin Lange[2,4], Alex Ingerman[2], Alison Lentz[2], Reed Enger[2], Andrew Barakat[2], Victoria Krakovna[1], John Oliver Siy[2], Zeb Kurth-Nelson[1], Amanda McCroskery[2], Vijay Bolina[1], Harry Law[1], Murray Shanahan[1], Lize Alberts[2,5,6], Borja Balle[1], Sarah de Haas[2], Yetunde Ibitoye[2], Allan Dafoe[1], Beth Goldberg[3], Sébastien Krier[1], Alexander Reese[2], Sims Witherspoon[1], Will Hawkins[1], Maribeth Rauh[1], Don Wallace[1], Matija Franklin[7], Josh A. Goldstein[8], Joel Lehman[9], Michael Klenk[10], Shannon Vallor[11], Courtney Biles[1], Meredith Ringel Morris[1], Helen King[1], Blaise Agüera y Arcas[2], William Isaac[1] and James Manyika[2]

[*]Equal contributions, [1]Google DeepMind, [2]Google Research, [3]Jigsaw, [4]Ludwig-Maximilians-Universität München, [5]University of Oxford, [6]Stellenbosch University, [7]University College London, [8]Center for Security and Emerging Technology, [9]Independent, [10]Delft University of Technology, [11]University of Edinburgh

Ethics of AI Assistants

- Value Alignment, Safety, and Misuse
  - Value Alignment
  - Well-Being
  - Safety
  - Malicious Uses

- Human-Assistant Interaction
  - Influence
  - Anthropomorphism
  - Appropriate Relationships
  - Trust
  - Privacy

- Assistants and Society
  - Cooperation
  - Access and Opportunity
  - Misinformation
  - Economic Impact
  - Environmental Impact
  - Evaluation

- Conclusions
  - Opportunities
  - Risks
  - Recommendations

Anthropomorphism – Human-like

COLUMBIA UNIVERSITY

- Human-Like physical features promote feelings of
  - Likability
  - Trust
  - Affinity

- People tend to attribute greater intentionality and intelligence to robot partners when their appearance was anthropomorphic than when robots appeared more mechanical.

Risk of Anthropomorphism

- If anthropomorphic design choices are not aligned with expectations users have of robotic interaction partners, designers run the risk of alienating audiences and fostering unfavorable impressions of robots.

- Humans experience extreme aversion to robots that appear human-like (the so-called ‘uncanny valley’) or perceive capable androids as threatening.

Risk of Anthropomorphism

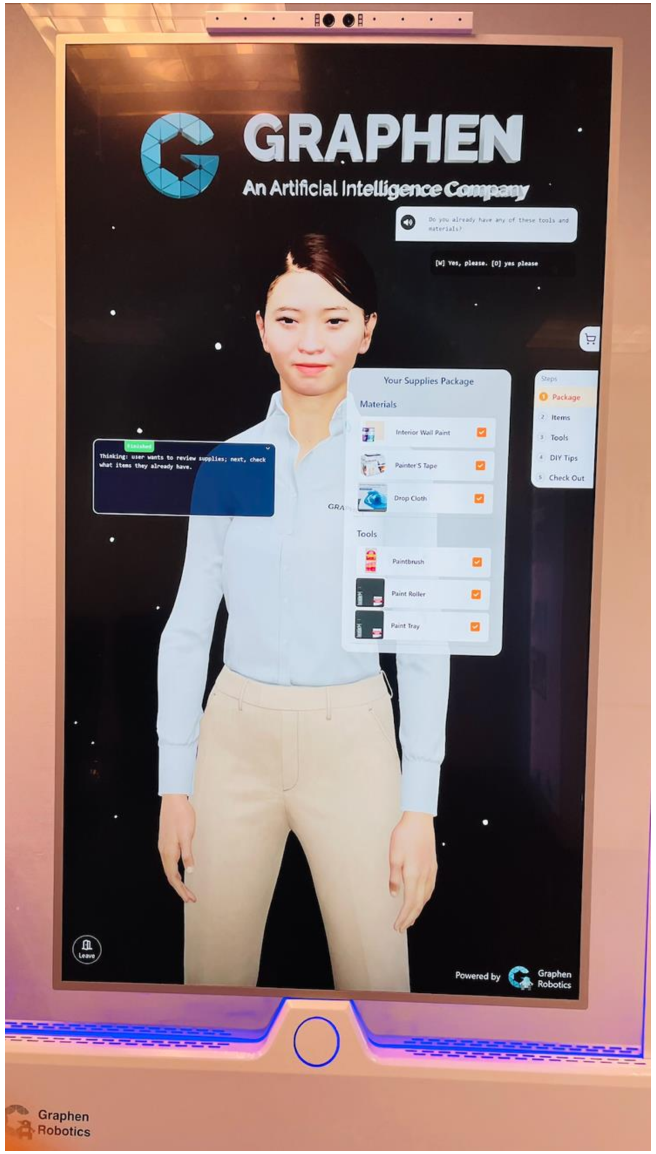
- Users may incorporate politeness conventions that are appropriate in use with other humans, but superfluous / rude when applied to exchanges with non-sentient AI.

Mechanisms that may enable harm

- Trust and Emotional Attachment.
- User trust has always been an aspirational end goal of building safe technology, be it robots or autonomous vehicles.
- Emotional attachment on the user’s behalf endows AI – and by extension, its creators – with considerable influence over a user’s thoughts, beliefs, emotions and psychological state.
- A being is considered human because it is human in essence, and no amount of resemblance and imitation can permit a non-human entity to encroach upon this categorization.

Anthropomorphism – Going Forward

- Trust
- Transparency
- Sound Design
- Redefine boundaries between “human” and “other”.



Columbia University

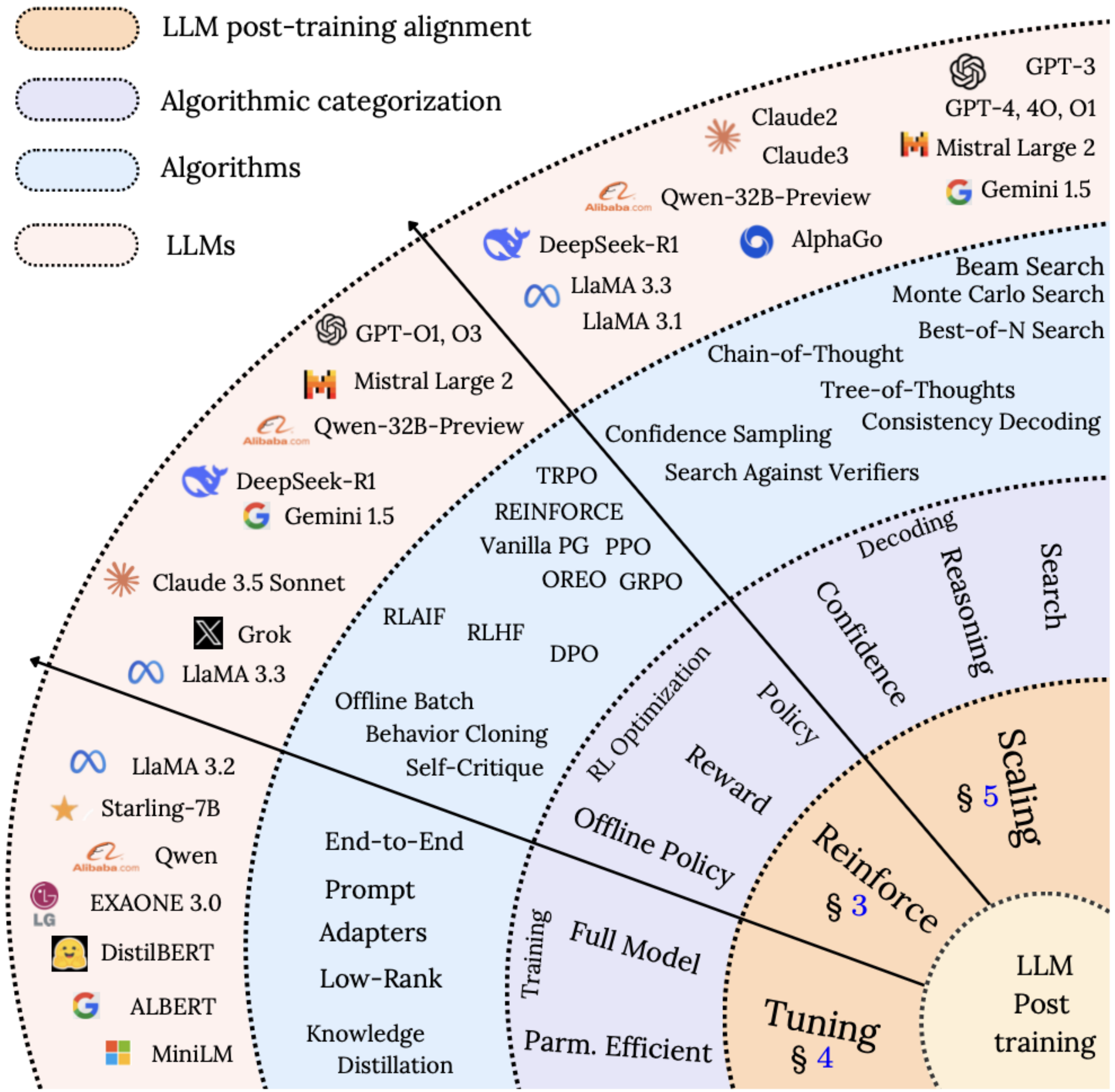
LLM Post-Training: A Deep Dive into Reasoning Large Language Models

Komal Kumar*, Tajamul Ashraf*, Omkar Thawakar, Rao Muhammad Anwer, Hisham Cholakkal, Mubarak Shah, Ming-Hsuan Yang, Phillip H.S. Torr, Fahad Shahbaz Khan, Salman Khan

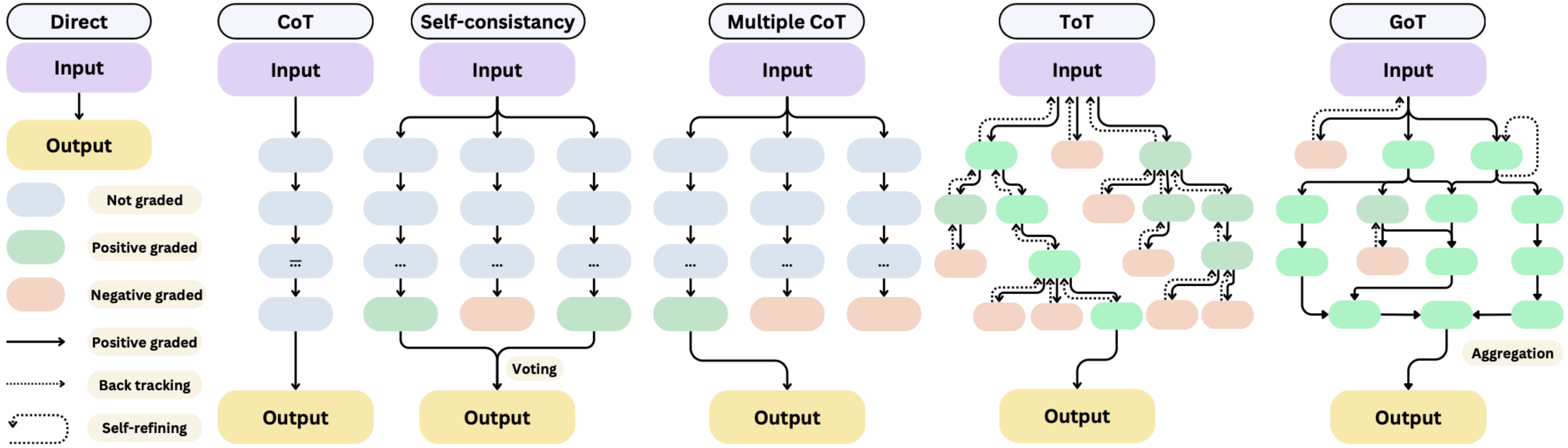
Abstract—Large Language Models (LLMs) have transformed the natural language processing landscape and brought to life diverse applications. Pretraining on vast web-scale data has laid the foundation for these models, yet the research community is now increasingly shifting focus toward post-training techniques to achieve further breakthroughs. While pretraining provides a broad linguistic foundation, post-training methods enable LLMs to refine their knowledge, improve reasoning, enhance factual accuracy, and align more effectively with user intents and ethical considerations. Fine-tuning, reinforcement learning, and test-time scaling have emerged as critical strategies for optimizing LLMs performance, ensuring robustness, and improving adaptability across various real-world tasks. This survey provides a systematic exploration of post-training methodologies, analyzing their role in refining LLMs beyond pretraining, addressing key challenges such as catastrophic forgetting, reward hacking, and inference-time trade-offs. We highlight emerging directions in model alignment, scalable adaptation, and inference-time reasoning, and outline future research directions. We also provide a public repository to continually track developments in this fast-evolving field: https://github.com/mbzuai-oryx/Awesome-LLM-Post-training.

3/24/2025

Reference for Post-Training LLMs



Reference for Post-Training LLMs



Reference for Post-Training LLMs

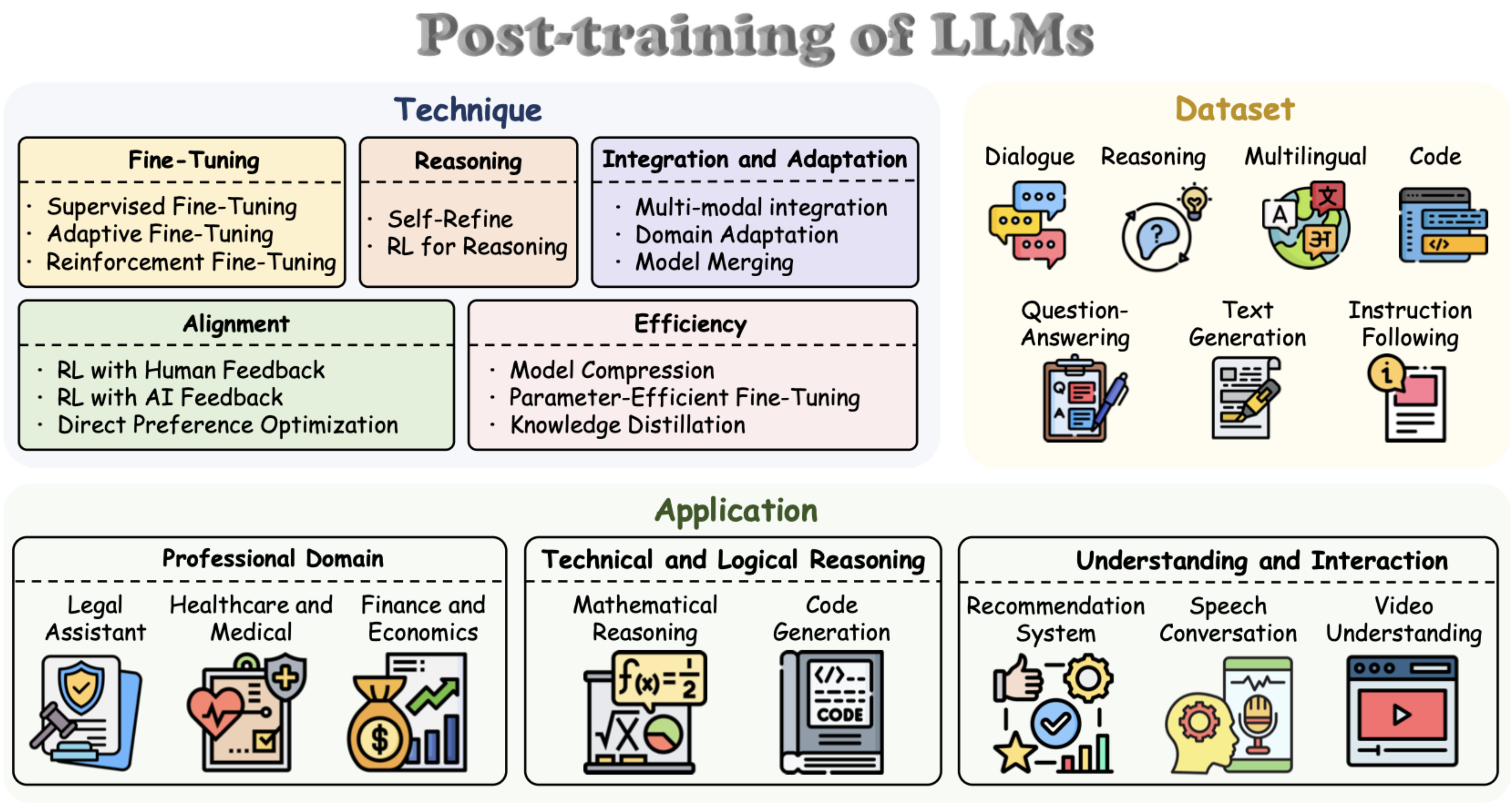
A Survey on Post-training of Large Language Models

Guiyao Tie[1†] Zeli Zhao[1] Dingjie Song[2] Fuyang Wei[3] Rong Zhou[2] Yurou Dai[2]
Wen Yin[1] Zhejian Yang[4] Jiangyue Yan[5] Yao Su[6] Zhenhan Dai[1] Yifeng Xie[1]
Yihan Cao[7] Lichao Sun[2] Pan Zhou[1] Lifang He[2] Hechang Chen[4] Yu Zhang[5]
Qingsong Wen[8] Tianming Liu[9] Neil Zhenqiang Gong[10] Jiliang Tang[11]
Caiming Xiong[12] Heng Ji[13] Philip S. Yu[14] Jianfeng Gao[15]

[1]Huazhong University of Science and Technology [2]Lehigh University
[3]The University of Hong Kong [4]Jilin University [5]Southern University of Science and Technology
[6]Worcester Polytechnic Institute [7]LinkedIn Corporation
[8]Squirrel Ai Learning [9]University of Georgia [10]Duke University
[11]Michigan State University [12]Salesforce Research [13]University of Illinois Urbana-Champaign
[14]University of Illinois at Chicago [15]Microsoft Research

3/8/2025

Reference for Post-Training LLMs



Google DeepMind

An Approach to Technical AGI Safety and Security

Rohin Shah[1], Alex Irpan[*,1], Alexander Matt Turner[*,1], Anna Wang[*,1], Arthur Conmy[*,1], David Lindner[*,1], Jonah Brown-Cohen[*,1], Lewis Ho[*,1], Neel Nanda[*,1], Raluca Ada Popa[*,1], Rishub Jain[*,1], Rory Greig[*,1], Samuel Albanie[*,1], Scott Emmons[*,1], Sebastian Farquhar[*,1], Sébastien Krier[*,1], Senthooran Rajamanoharan[*,1], Sophie Bridgers[*,1], Tobi Ijitoye[*,1], Tom Everitt[*,1], Victoria Krakovna[*,1], Vikrant Varma[*,1], Vladimir Mikulik[*,2], Zachary Kenton[*,1], Dave Orr[1], Shane Legg[1], Noah Goodman[1], Allan Dafoe[1], Four Flynn[1] and Anca Dragan[1]
[1]Google DeepMind, [2]Work done while at Google DeepMind, [*]Core contributor, alphabetical order

Artificial General Intelligence (AGI) promises transformative benefits but also presents significant risks. We develop an approach to address the risk of harms consequential enough to significantly harm humanity. We identify four areas of risk: misuse, misalignment, mistakes, and structural risks. Of these, we focus on technical approaches to misuse and misalignment. For misuse, our strategy aims to prevent threat actors from accessing dangerous capabilities, by proactively identifying dangerous capabilities, and implementing robust security, access restrictions, monitoring, and model safety mitigations. To address misalignment, we outline two lines of defense. First, model-level mitigations such as amplified oversight and robust training can help to build an aligned model. Second, system-level security measures such as monitoring and access control can mitigate harm even if the model is misaligned. Techniques from interpretability, uncertainty estimation, and safer design patterns can enhance the effectiveness of these mitigations. Finally, we briefly outline how these ingredients could be combined to produce safety cases for AGI systems.

4/2/2025


GRAPHEN
www.graphen.ai
Advancing AI for Well-Being of the Mankind