

# E6893 Big Data Analytics Lecture 12:

## *Big Data and AI Applications in Finance Industry*

Ching-Yung Lin, Ph.D.

Adjunct Professor, Dept. of Electrical Engineering and Computer Science



December 1st, 2023

# Before

# After







- **How does FINRA analyze ~50B events per day TODAY?** - *Build a graph of market order events from multiple sources* [[ref](#)]

# Example: Help reduce AML cost

## Tremendous Compliance Costs & Workload

48 Days

onboarding new customers

30

Minutes to clear a generated alert

X 20,000

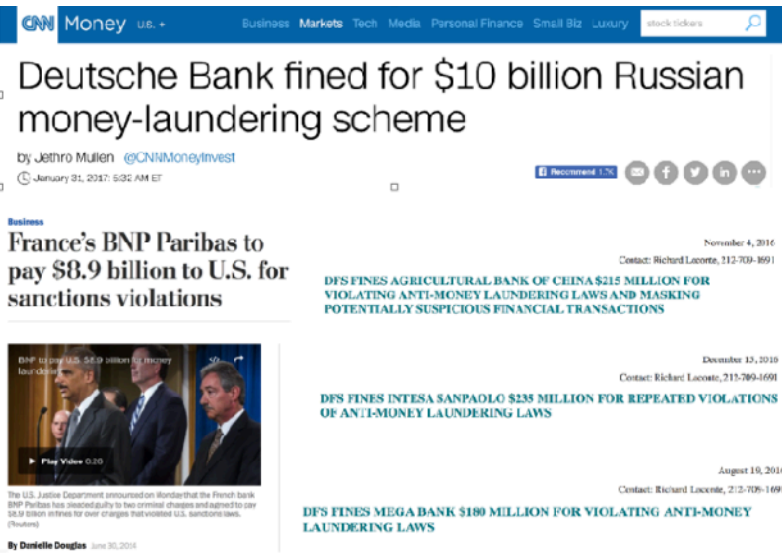
Alerts per day

=

10,000

= 1350 Analysts

Total hours lost per day



### US Federal Reserve for violating anti-money laundering act

All the three branches in New York, Chicago and Silicon Valley were fined

Like 32 Share Tweet G+ Share

By Taiwan News, Staff Writer

2018/01/18 16:58

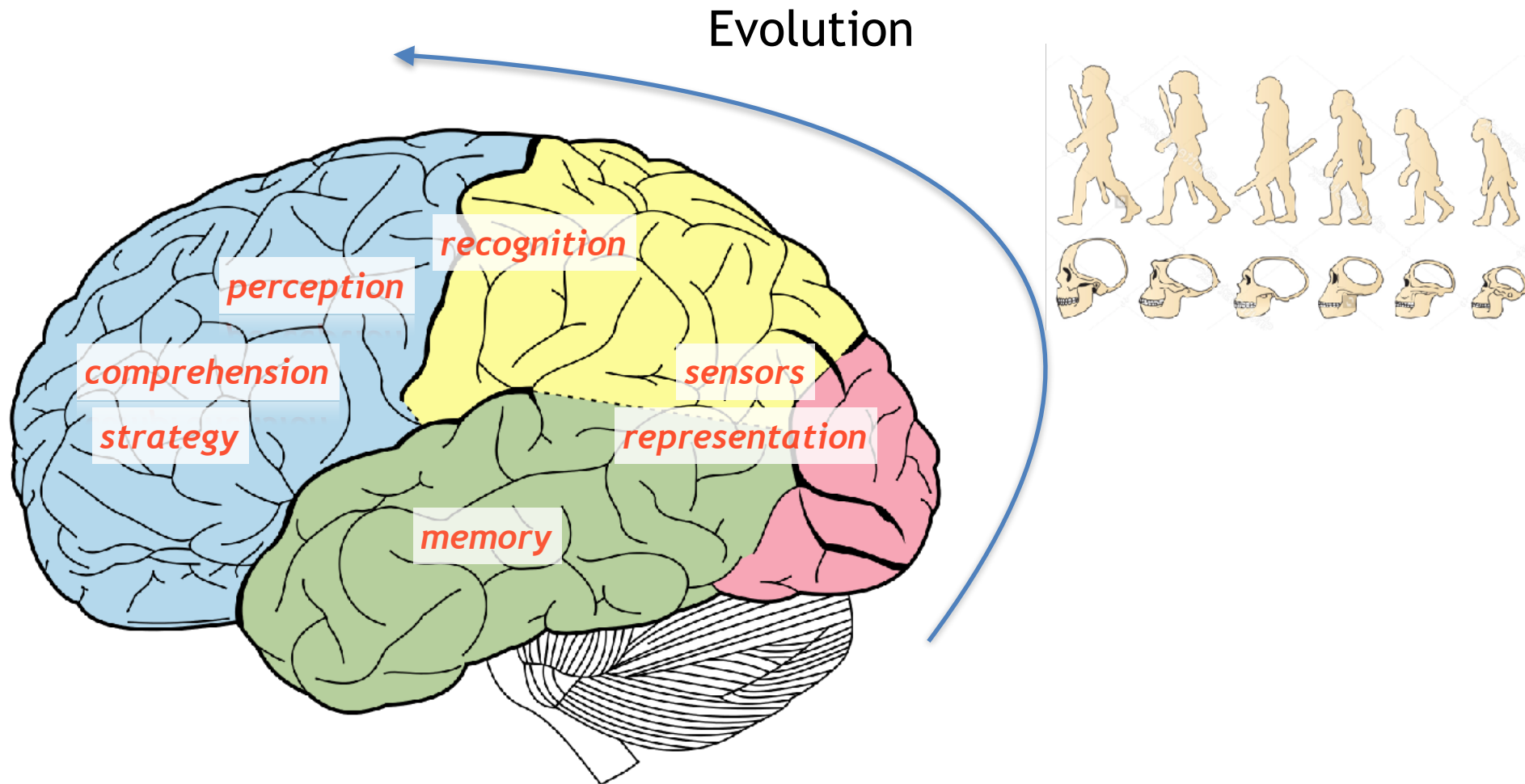
THU JAN 4, 2018 / 1:38 PM EST

Citibank fined \$70 million for anti-money laundering compliance shortcomings





- How did journalists uncover the Swiss Leak scandal in 2014 and also Panama Papers in 2016? -- Using graph database to uncover information thousands of accounts in more than 20 countries with links through millions of files [\[ref\]](#)



# Future of AI ==> Full Function Brain Capability

Most of existing “AI” technology is only a key fundamental component.

- **Machine Cognition:**
  - Robot Cognition Tools
  - Feeling
  - Robot-Human Interaction

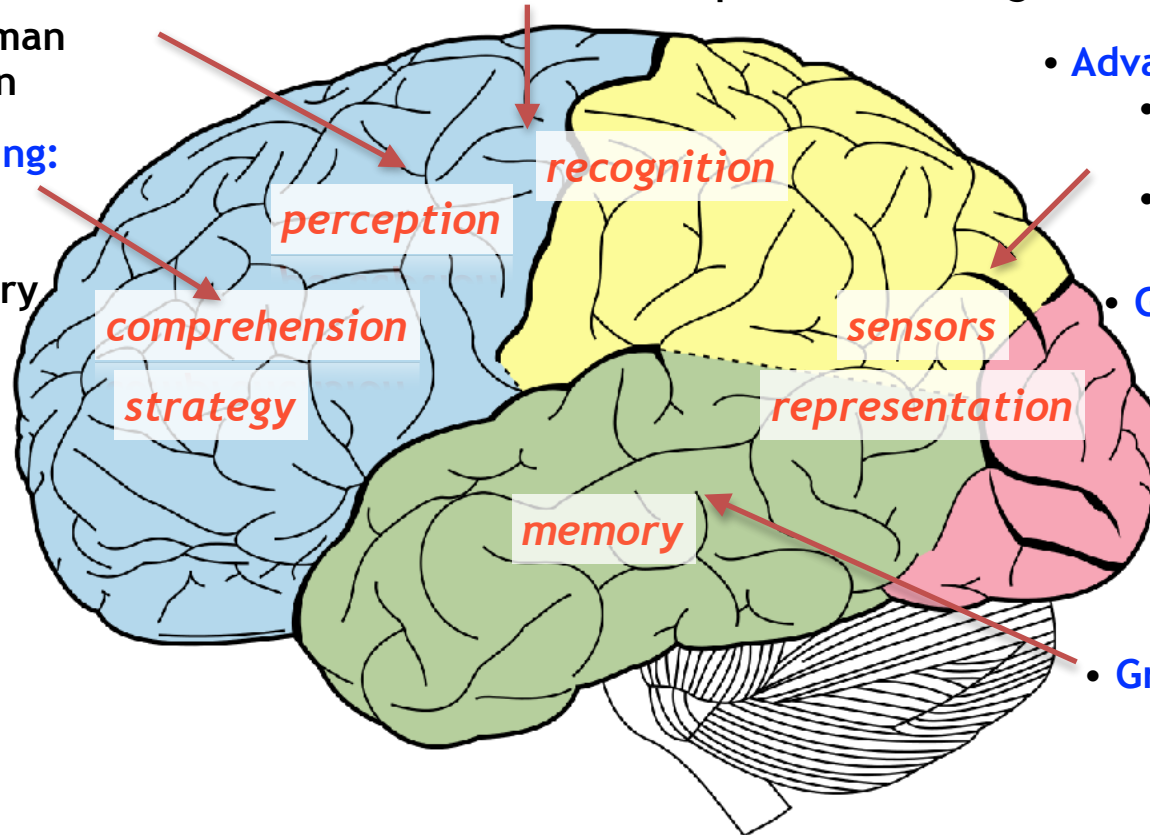
- **Machine Reasoning:**
  - Bayesian Networks
  - Game Theory Tools

- **Machine Learning:**
  - ML and Deep Learning
  - Autonomous Imperfect Learning

- **Advanced Visualization:**
  - Dynamic and Interactive Viz.
  - Big Data Viz.

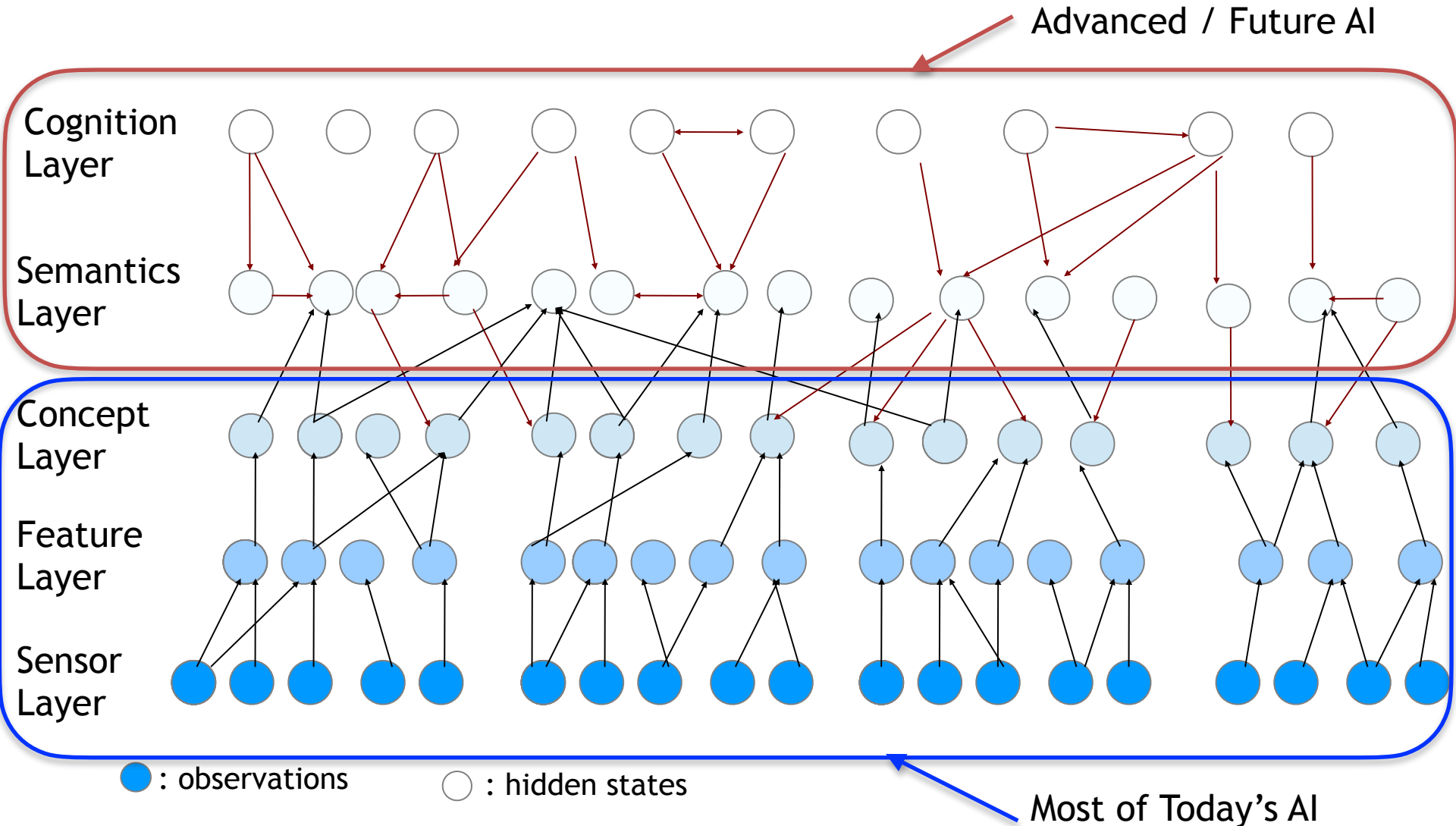
- **Graph Analytics:**
  - Network Analysis
  - Flow Prediction

- **Graph Database:**
  - Distributed Native Database





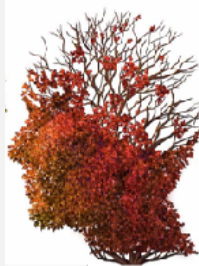
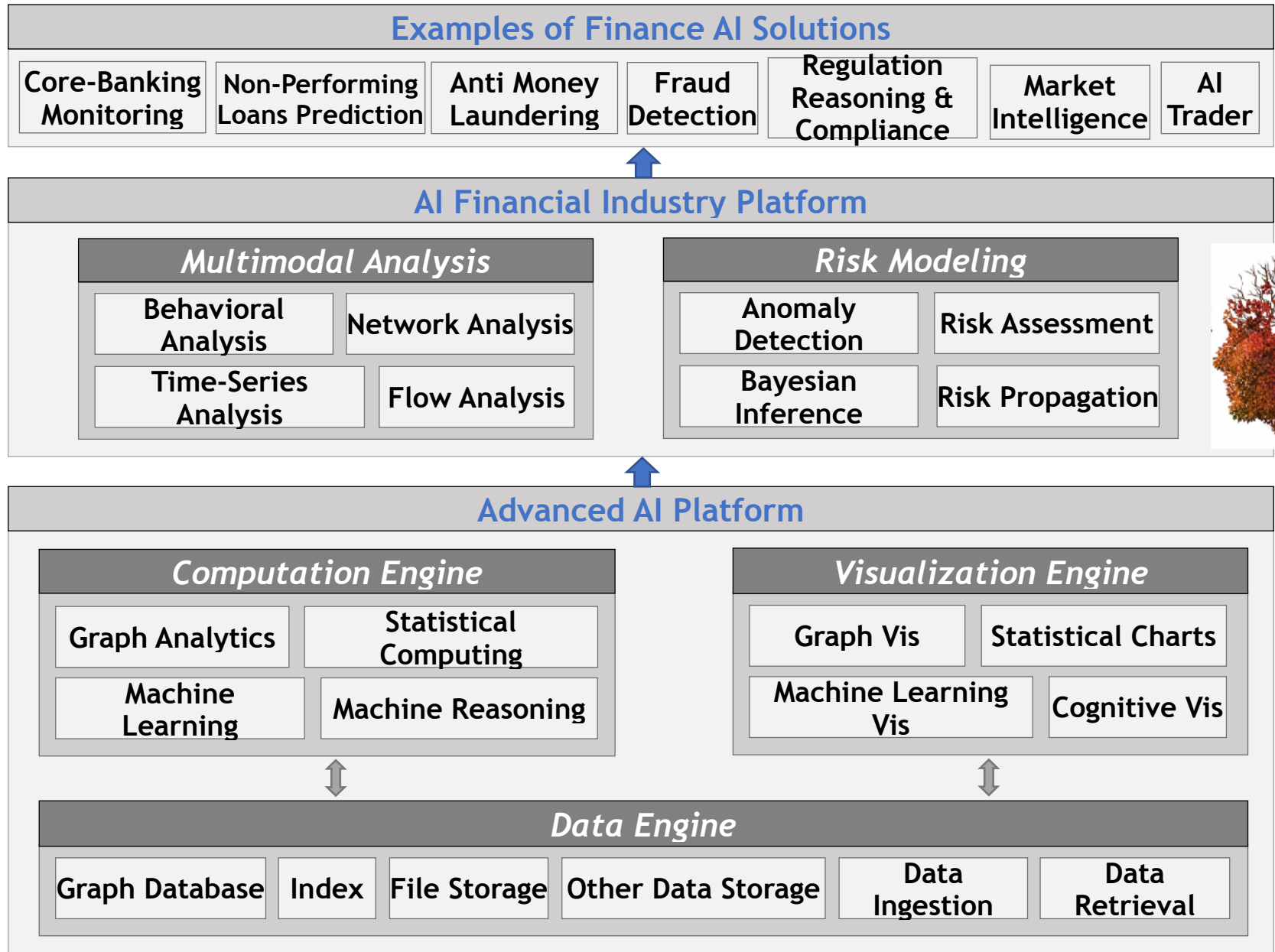
# Layers of Artificial Intelligence





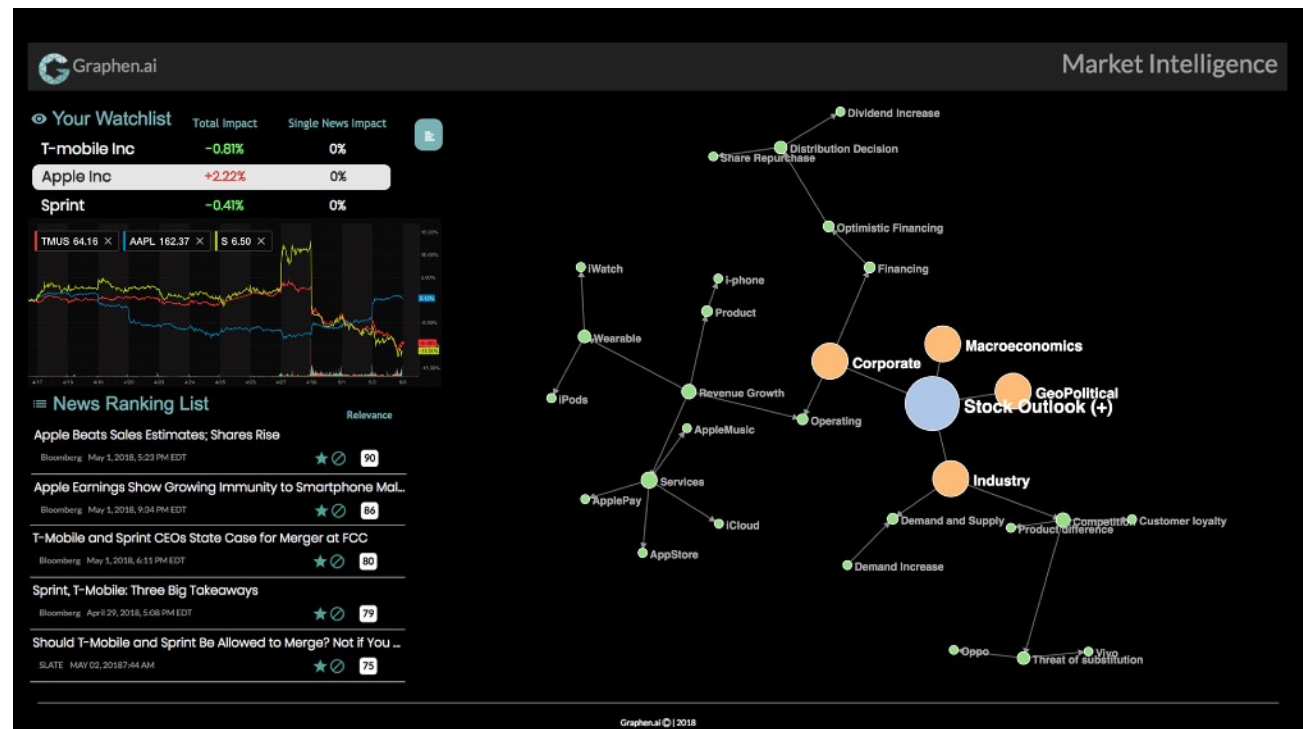
**AI Makes Safer, more Intelligent,  
and more Efficient Banks**

# Example of AI Finance Platform (Graphen Ardi Platform)





- Significantly improved Non-Performing-Loan accuracy rate in one of the world's largest banks (from ~20% prediction accuracy to ~60% accuracy).
- Advanced Anti-Money Laundering for banks — capable of predicting unknown unknowns.
- Detecting Fraud from Real-Time on Transactions in one of the world's largest transaction platform — on the scale of billions.
- Analyzing relationship data for an European bank.
- Cyber and Physical Security for another European bank.






# **Short Introduction of AI Loan Risk Prediction Solution**

Utilize AI Platform to predict NPL via:

- Mining relationship of customers
- Analyzing all kinds of network topological structure
- Understanding the entities
- Predicting the bad loan flows



**Ardi**  
Graph DB and  
Graph Analytics



Machine Learning Classifiers

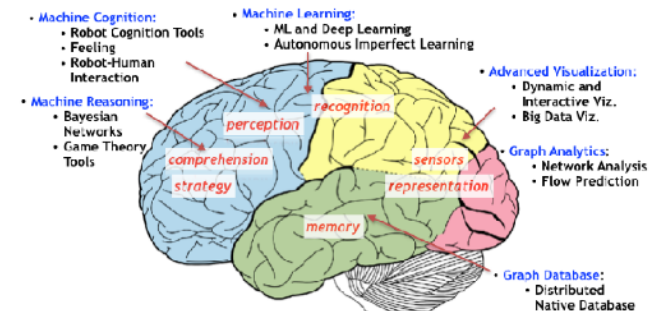
Dynamic Flow Learning

Deep Learning

**Ardi Machine Learning Tools**

**Ardi Machine Reasoning Tools**

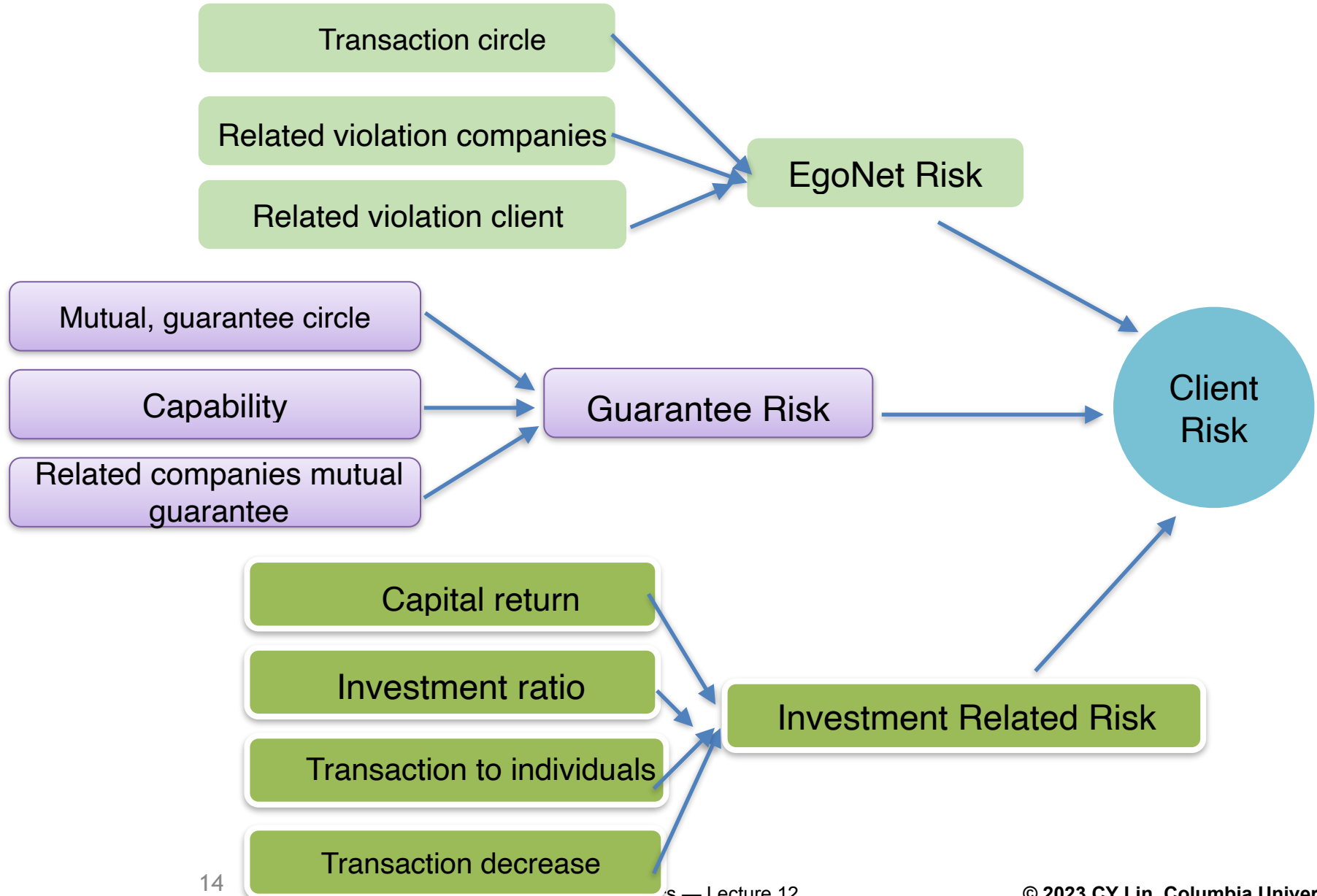
Cognitive Reasoning



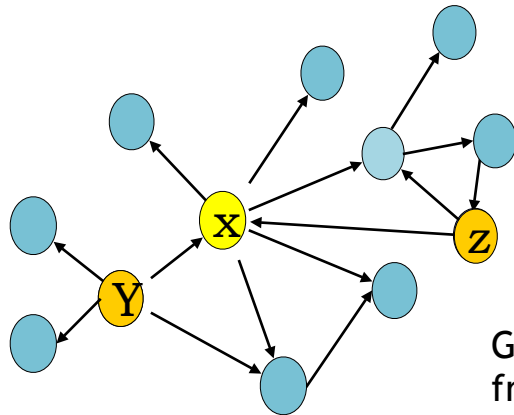
*==> Improving the accuracy from < 20% to near 60%*



# Debt Performing Analysis to determine the Risk Factors

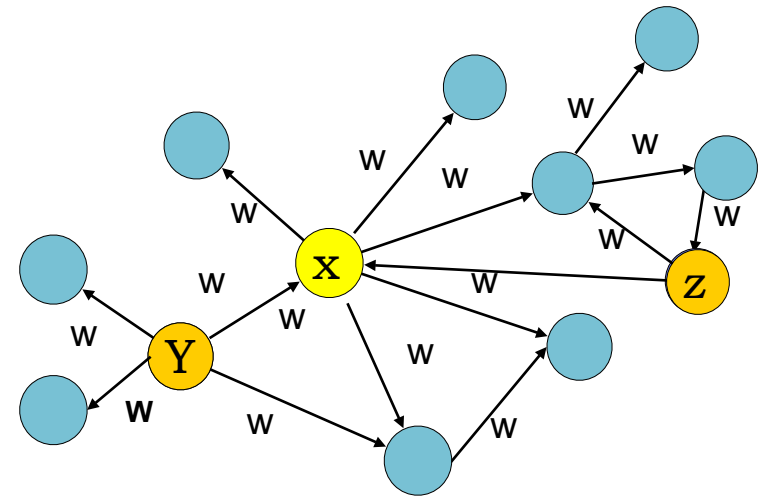


# Graph Analysis to Predict Risk



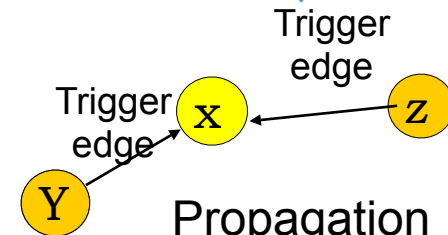
Directional Graph

Graph feature extraction  
from multiple relationships



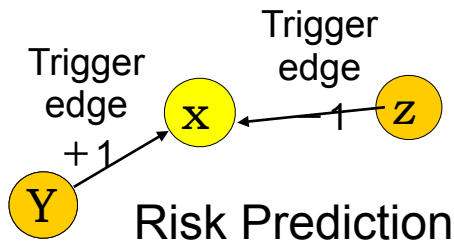
Weighted Directional Graph

Risk Propagation  
Edge Generation



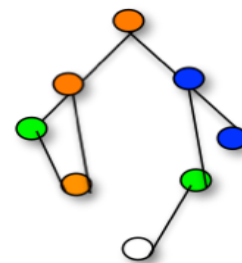
Propagation  
Directional  
Graph

Rick prediction based on  
machine learning and  
dynamic flow.

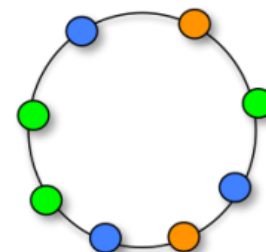


Risk Prediction

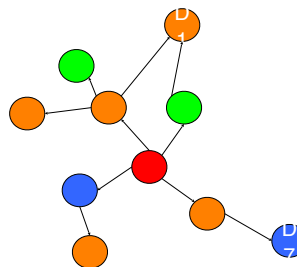
- Breadth First Search (BFS)
- Centrality Computation
- (Strong) Connected Component
- Cycle Detection
- Ego Net
- Pagerank
- Shortest Path (Dijkstra)



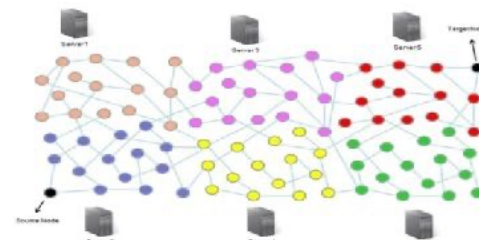
**Connected component**



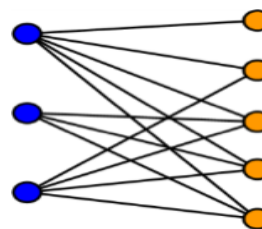
**Cycle detection**



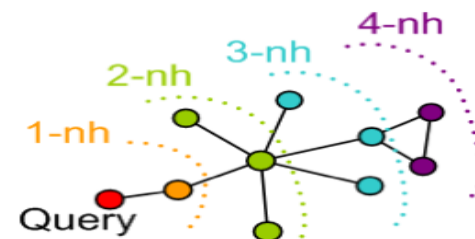
**Egonet**



**Shortest paths**



**Collaborative filtering**



**K-neighborhood**





# **Introduction of AI Anti-Money-Monitoring Solution**

## Increased Regulatory Expectations & Enforcement of Current Regulations

- AML laws and regulations keep evolving and become increasingly daunting. AML programs and IT systems need to be updated to ensure effectiveness and efficiency.
- Regulated institutions have to address recent regulatory changes to ensure that their transaction monitoring and filtering programs are designed to comply with regulatory standards and expectations.

According to a survey conducted by Dow Jones and ACAMS on 812 respondents in 2016, 60% of respondents cited this as the greatest AML compliance challenge. Over 75% of respondents cited FinCEN's proposed Beneficial Owner Rules as a contributor to **increased workloads and shortage of trained staff**, followed by FATCA, other tax evasion legislation and Fourth EU Anti-Money Laundering Directive.

According to ACAMS 2017 survey, financial institutions in U.S. have had to perform sweeping overhauls of their customer screening, monitoring and reporting processes courtesy of the U.S. Treasury's Office of Foreign Asset Control (OFAC). Changing OFAC priorities have significantly impacted operations for 53 percent of survey respondents who regularly engage in sanctions screening.

On Dec. 21<sup>st</sup> 2016, the European Commission proposed a series of legislative proposals, hoping to strengthen the EU's legal framework for anti-money laundering, controlling illegal cash flow, and freezing confiscation of illegal assets, thus strengthening the EU's efforts to fight against terrorism and organized crime.

On Jul. 25 2016, the Monetary Authority of Singapore (MAS) stated that it will step up anti-money laundering controls and take prompt actions against the banking industry. Previous investigations have revealed that several financial institutions based in Singapore involved funds related to the 1MDB scandal.

# 2017 AML Penalties -- I

| Date      | Punished Agency                               | Regulatory Agency                                | CCY | Amount      |
|-----------|---|--|-----|-------------|
| 13-Jan-17 | Toronto Dominion Bank of Canada               | U.S. Treasury Department Overseas Control Office | USD | 520,000     |
| 19-Jan-17 | Western Union Financial Services              | U.S. Financial Crime Enforcement Agency          | USD | 184,000,000 |
| 30-Jan-17 | German Deutsche Bank                          | New York State Financial Services Bureau         | USD | 425,000,000 |
| 30-Jan-17 | German Deutsche Bank                          | UK Financial Conduct Authority                   | GBP | 163,000,000 |
| 9-Feb-17  | Japan Mitsubishi Tokyo Union Bank             | British Prudential Regulation Authority          | GBP | 17,850,000  |
| 16-Feb-17 | Tabcorp                                       | Australian Trading Report and Analysis Center    | AUD | 45,000,000  |
| 17-Feb-17 | Australian Trading Report and Analysis Center | Florence, Italy Prosecutor                       | EUR | 600,000     |
| 27-Feb-17 | California Merchants Bank                     | U.S. Financial Crime Enforcement Agency          | USD | 7,000,000   |
| 7-Mar-17  | Guangdong Securities                          | Hong Kong Securities Regulatory Commission       | HKD | 3,000,000   |
| 14-Mar-17 | Sino-Thai International Securities            | Hong Kong Securities Regulatory Commission       | HKD | 2,600,000   |
| 5-Apr-17  | Guoyuan Securities Broker (Hong Kong)         | Hong Kong Securities Regulatory Commission       | HKD | 4,500,000   |
| 11-Apr-17 | British Careers Bank                          | Hong Kong Monetary Authority                     | HKD | 7,000,000   |
| 26-Apr-17 | Irish Union Bank                              | Irish Central Bank                               | EUR | 2,200,000   |
| 22-May-17 | National Bank of Mexico, United States Branch | United States Department of Justice              | USD | 97,440,000  |
| 30-May-17 | German Deutsche Bank                          | Federal Reserve Board                            | USD | 41,000,000  |
| 30-May-17 | Irish Bank                                    | Irish Central Bank                               | DUR | 3,150,000   |

# 2017 AML Penalties - II

| Date      | Punished Agency                      | Regulatory Agency   | CCY | Amount      |
|-----------|--------------------------------------|---|-----|-------------|
| 31-May-17 | Credit Suisse Bank                   | Singapore Financial Regulatory Commission                     | SGD | 700,000     |
| 31-May-17 | Singapore UOB Bank                   | Singapore Financial Regulatory Commission                     | SGD | 900,000     |
| 3-Jun-17  | BNP Paribas                          | French Prudential Supervision Association                     | EUR | 10,000,000  |
| 16-Jun-17 | Bank of China Milan Branch           | Italian tax authorities and the Ministry of Economic Affairs  | EUR | 20,000,000  |
| 22-Jun-17 | Edmund Rothschild Group, Switzerland | Luxembourg Financial Supervision Commission                   | EUR | 9,000,000   |
| 6-Jul-17  | Latvian Rito Bank                    | A court in Paris  | EUR | 80,000,000  |
| 27-Jul-17 | Bitcoin Trading Platform btc-e       | U.S. Financial Crime Enforcement Agency                       | USD | 110,000,000 |
| 3-Aug-17  | Australian Commonwealth Bank         | Australian Trading Report and Analysis Center                 | AUD | 18,000,000  |
| 7-Sep-17  | Habib Bank of Pakistan               | New York State Financial Services Agency                      | USD | 225,000,000 |
| 29-Sep-17 | Real Madrid multibank                | Panama Banking Authority                                      | USD | 300,000     |
| 1-Nov-17  | Lone Star Bank                       | U.S. Financial Crime Enforcement Agency                       | USD | 2,000,000   |
| 27-Nov-17 | Italy Union Bank of Sao Paulo        | Irish Central Bank  | EUR | 1,000,000   |
| 21-Dec-17 | Merrill Lynch                        | U.S. Securities and Exchange Commission                       | USD | 13,000,000  |
| 21-Dec-17 | Korea Agricultural Association Bank  | New York State Financial Services Agency                      | USD | 11,000,000  |
| 21-Dec-17 | Danske Bank                          | Danish Severe Economic and International Crime Control Agency | DKK | 12,500,000  |

Source: <http://www.yangqiu.cn/HT-AML/4007141.html>



## Legacy/Outdated Systems & Inaccurate Results

- Financial institutions are saddled with legacy AML compliance systems that were built piecemeal and can no longer meet current needs and regulatory expectations.
- The work flow involves multiple separate IT systems and databases and thus requires significant manual work for data integration and retrieval.
- Previous AML systems have limited capability to uncover hidden relationships among customers and accounts.
- Rules are often created based on specific scenarios but don't consider different contexts of individuals, resulting in many **false positives** which make alert review labor-intensive and time consuming.

Average time to clear a generated alert requires 19 mins in 2016. For a bank reporting 10,000 alerts per day, it's a lost of 3,167 hours.

|   | 2013 | 2015 | 2016 |
|---|------|------|------|
| Cost issues                                 | 38%  | 40%  | 46%  |
| Too many false-positive alerts              | 48%  | 47%  | 45%  |
| Too many false-negative alerts              | 40%  | 40%  | 43%  |
| Poor customer service/roadmap               | 31%  | 33%  | 36%  |
| Consolidation of software across risk areas | 32%  | 38%  | 35%  |
| Enterprise technology consolidation         | 33%  | 34%  | 35%  |
| Inability to handle non-Latin script        | 10%  | 12%  | 11%  |

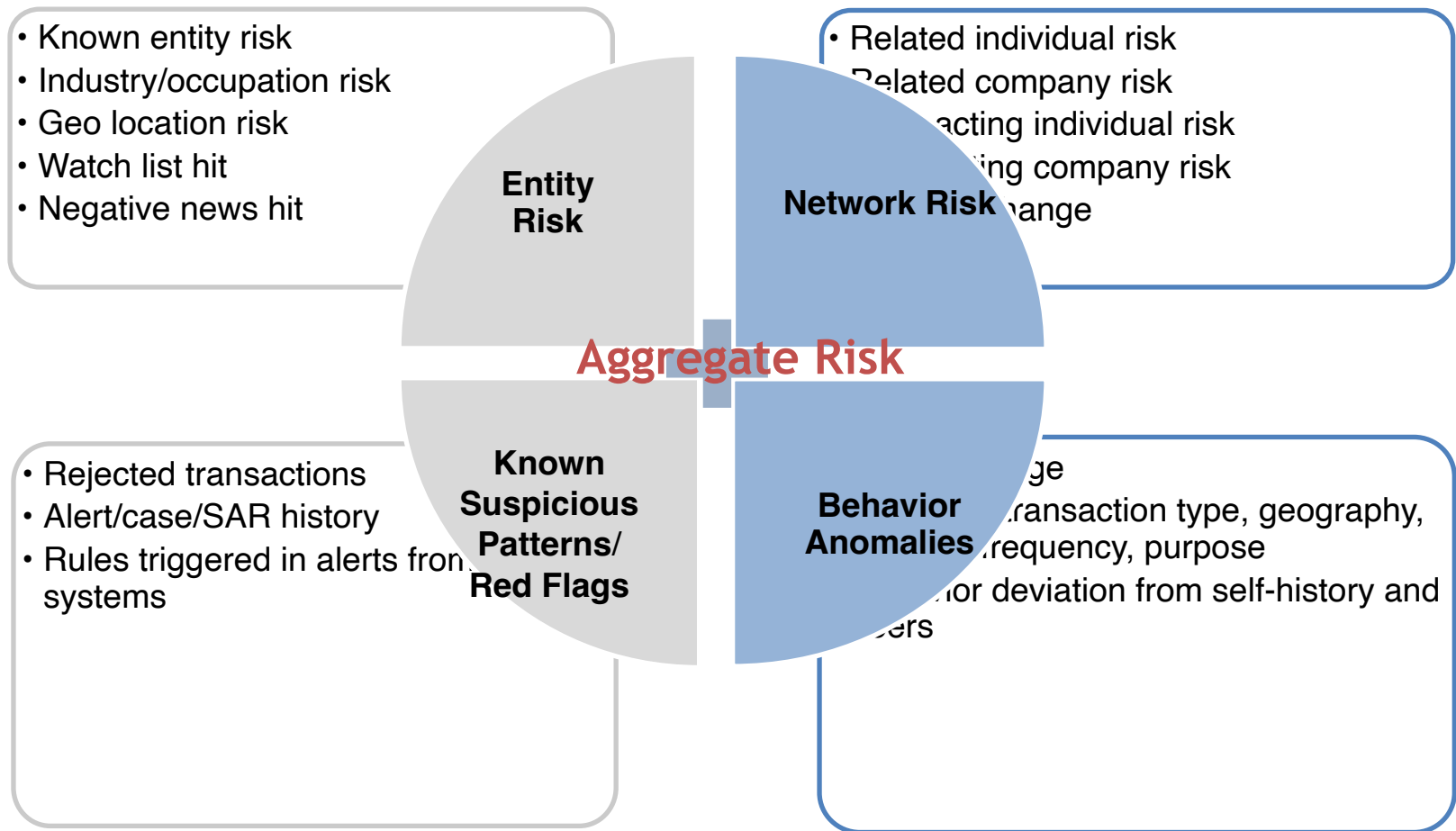
Source: Dow Jones & ACAMS 2016

## New Money Laundering Techniques – Unknown Unknowns

- Money launderers change techniques over time, and FATF has to update its recommendations every couple of years.
- The use of proxy servers and anonymizing software makes the third component of money laundering, integration, almost impossible to detect, as money can be transferred or withdrawn leaving little or no trace of an IP address.
- The use of the internet allows money launderers to easily avoid detection. The rise of online banking institutions, anonymous online payment services, peer-to-peer transfers using mobile phones and the use of virtual currencies such as Bitcoin have made detecting the illegal transfer of money even more difficult.
- Adding rules to cover newly discovered money laundering techniques and patterns requires a lot of manual work.
- Rule-based transaction screening system can only identify suspicious transactions with known patterns; however, there are always new money laundering techniques. It is very important for banks to catch the “**unknown unknowns** (things we don’t know we don’t know)”.

# How does Graphen AI Help AML?

- Automatically considers features from multiple aspects to more accurately assess the risk of each party.
- Automatically builds activity-based behavior models, analyzes every party's current behavior in the context of self-history and the behavior of peers, and identify behavior outliers.



## Traditional AML System

AI-Powered AML System

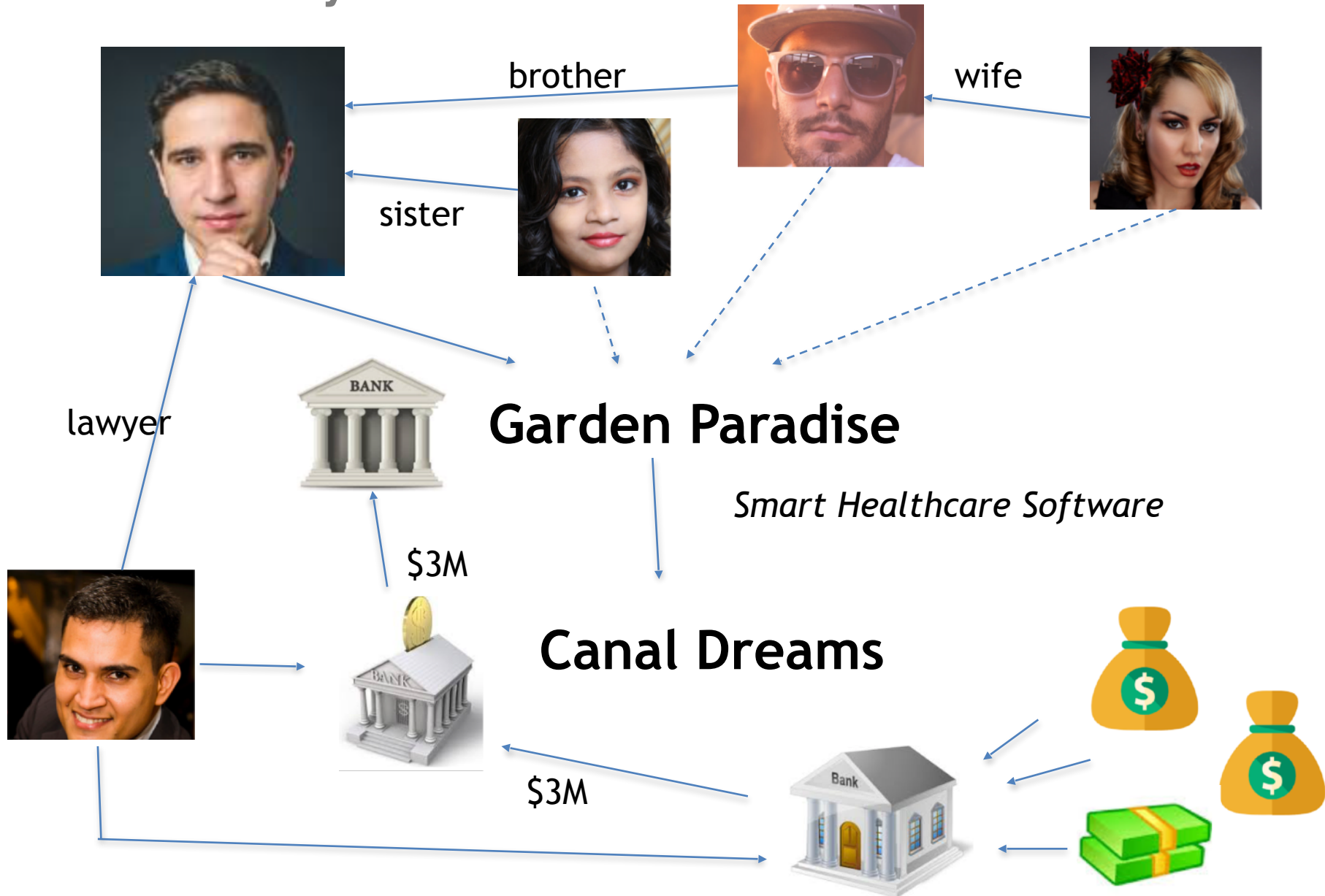
|                       |        | SAR  | No SAR  |
|-----------------------|--------|--|---|
| AI-Powered AML System | SAR    | <ul style="list-style-type: none"><li>Both systems agree on SAR.</li></ul>   | <ul style="list-style-type: none"><li>Finds “unknown unknowns” by detecting outliers and identifying abnormal patterns.</li><li>Reduces the risk of AML compliance failure.</li></ul> |
|                       | No SAR | <ul style="list-style-type: none"><li>Reduces “false positives” via automatic EDD and in-context analysis of accounts and parties.</li><li>Reduces manual workload to filter false alarms.</li><li>Improves efficiency with aggregate risk ranking and automatic retrieval of case-relevant data for investigation</li></ul> | <ul style="list-style-type: none"><li>Both systems agree on no SAR.</li></ul>   |



Jose Rose is a drug dealer from Central America living in New York. In order to send his money to US, he set up a Company – Garden Paradise in New Jersey. Its legal representative and managers are Mr. Rose and his relatives. He opened an account at a local C Bank in Jersey City. However, the company's book seemed to have failed to explain Mr. Rose's huge annual income. So he needed a money laundering plan, executed in five steps:

1. Over a few years, he slowly transported huge sums of money to Panama through various means to Bank A (via mail, shipping, entrained in various goods).
2. Mr. Rose flew to Panama to find a lawyer and set up a Company - Canal Dreams, with an account opened at Bank B in Panama.
3. Mr. Rose represents Garden Paradise to negotiate with Canal Dreams. They 'discussed' selling a service contract of Smart Healthcare software to Garden Paradise and finally signed a contract of 3 million US dollars.
4. At this point, Mr. Rose's secretly authorized lawyer transferred the money he has deposited in Bank A to Canal Dreams' account in Bank B.
5. Wait until the legal procedures were completed. The money was then transferred from Bank B account to Mr. Rose's C Bank account in New Jersey.

# Use Case Story

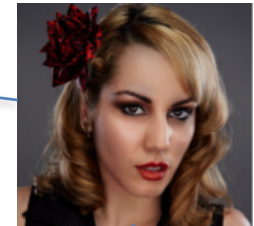




brother



wife



sister



AI: Who are their employees?

AI: What did they sell?  
Any other 'related' customers?

AI: Is this company capable of creating expensive software?

AI: Is it related to Healthcare business?

## Garden Paradise

*Smart Heal*

## Canal Dreams



\$3M



\$3M



AI: What are their relationships?  
What did they discuss in legal documents?



# User Interface Demo (I)

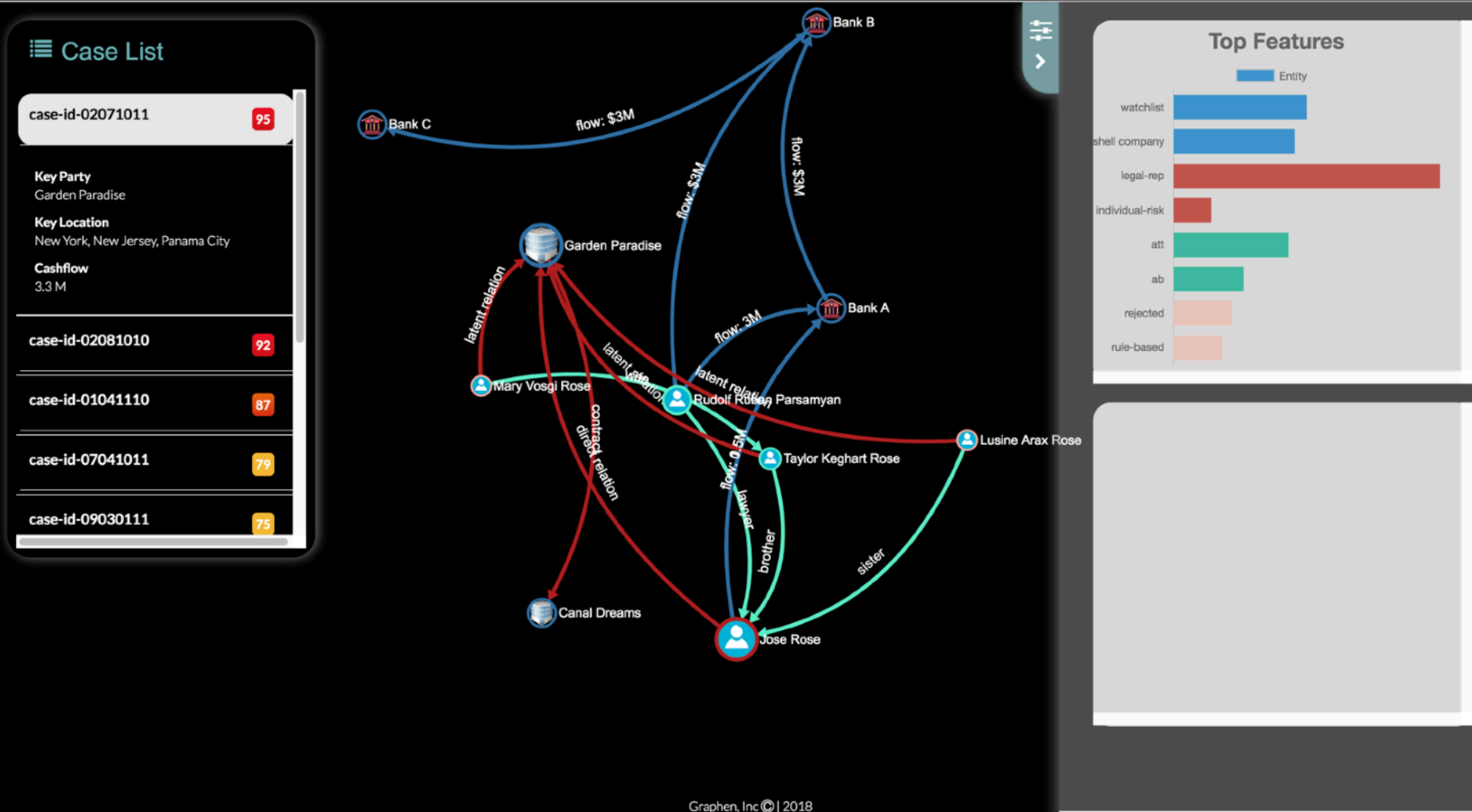


- Highlight the most suspicious case, including the involved entities,



# User Interface Demo (II)

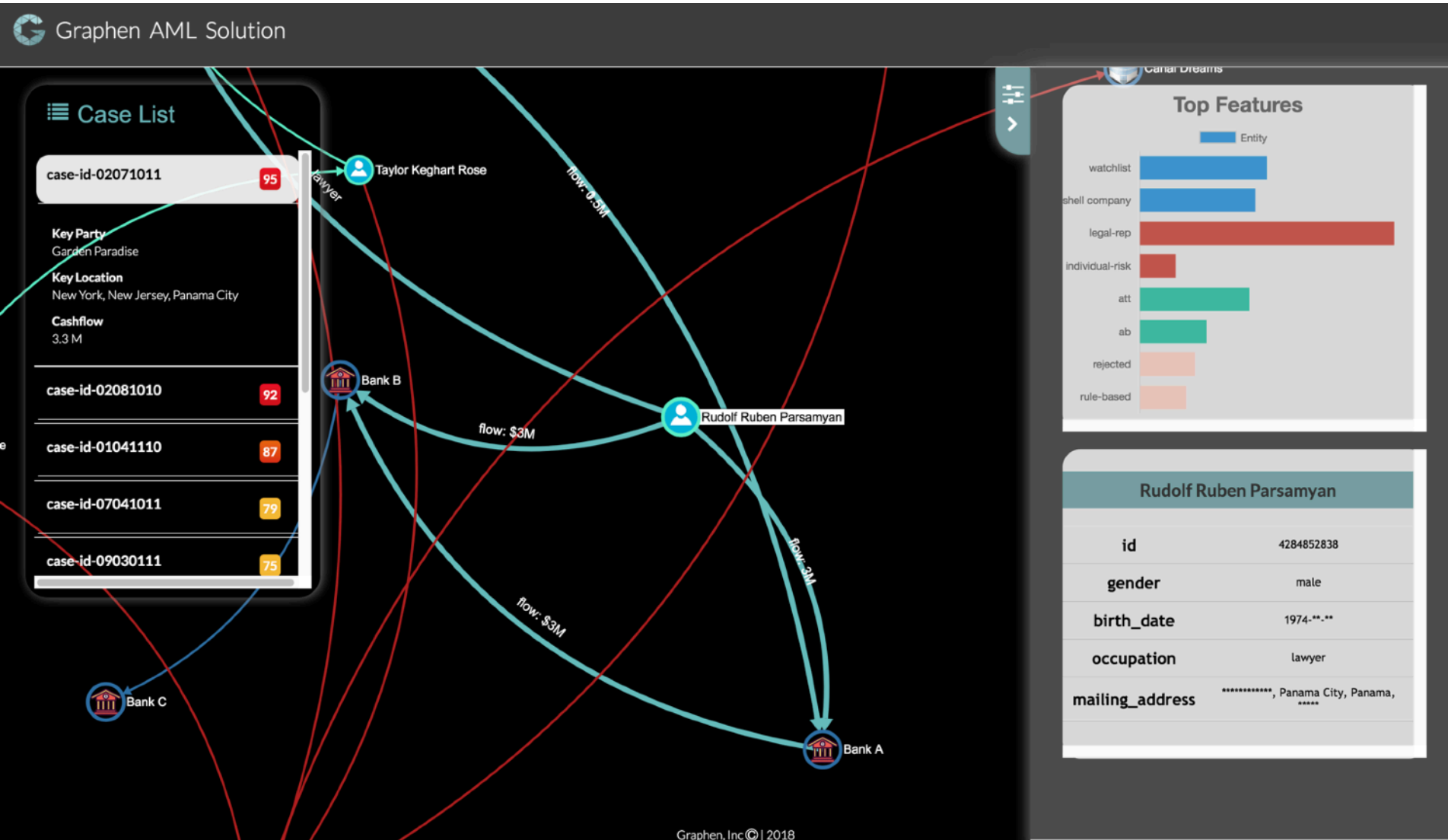
## Graphen AML Solution



Graphen, Inc. © 2018

- Show top anomaly features of the four categories: entity, network, pattern, and behavior. Highlight abnormal relationships in the graphs.

# User Interface Demo (III)



- Investigate suspicious individuals

## Example

A party has 5 international wire transfers of \$10K+ each time within last month.

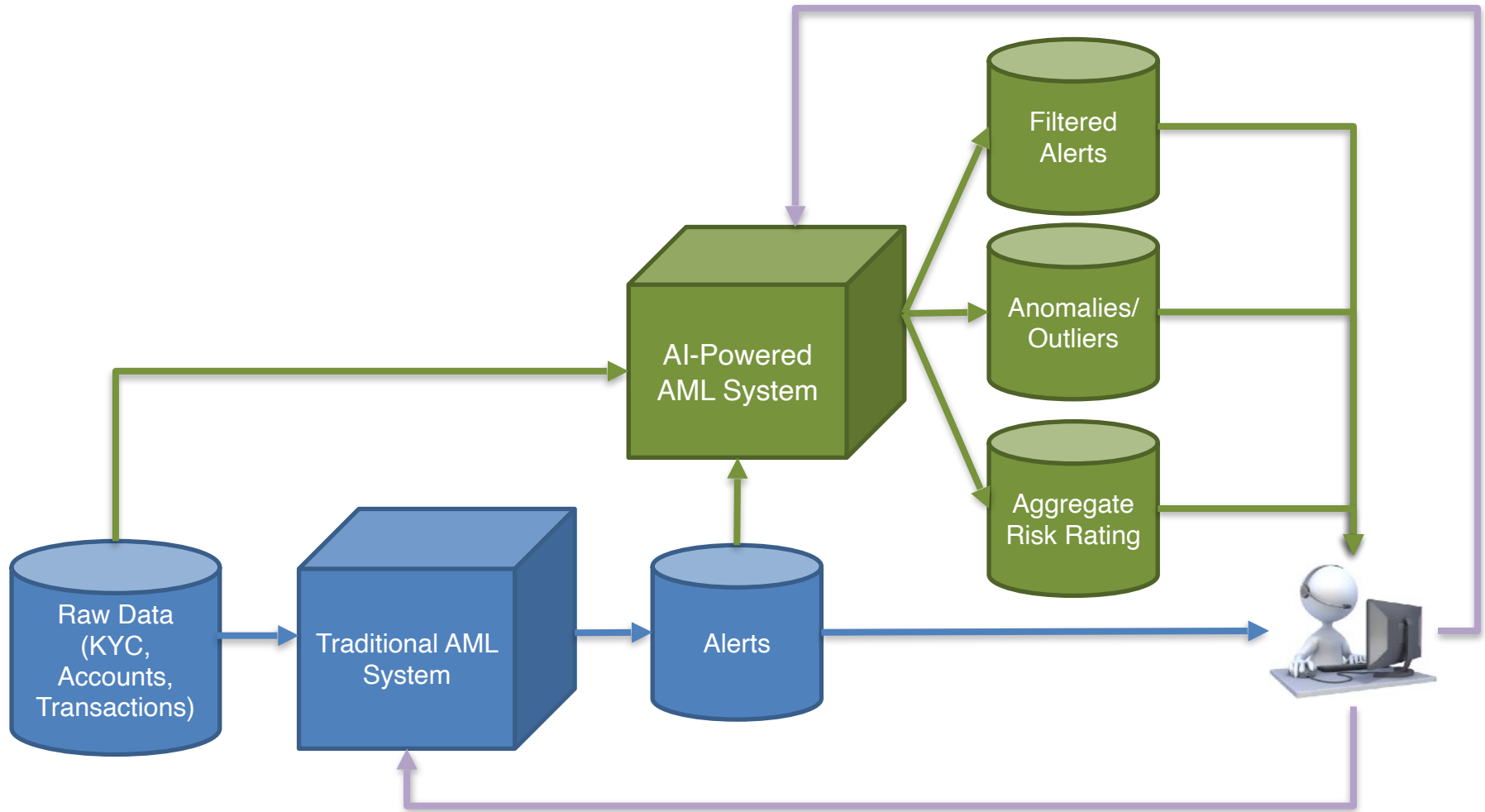
### Traditional Rule-Based AML

1. Matches the rule of “3+ international wire transfers of at least \$10K each within 30 days”.
2. Creates an alert.

### AI-Powered AML

1. Computes the answers to the following questions:
  - What is the occupation or business nature of this party?
  - How often did this party have international wire transfers before?
  - Who are the most frequent debit/credit counterparties of this party?
  - Has this party had transactions before with the counterparties of these wire transfers?
  - Does this party have other relationships (e.g. same account, address, phone, email) with these counterparties?
  - What are the case/SAR histories of this party and the counterparties?
  - Do peers of this party (e.g. individuals with the same occupation, or businesses of the same nature) often have wire transfers of similar frequency and amount?
  - ...
2. Determines the aggregate risk of this party based on all of the above answers.
  - Import/export company with frequent foreign trades → low risk (**false positive**).
  - A salon worker with small-amount of monthly activities and few wire transfers → high risk.

# Integrate AI-Powered AML with Current Workflow





## Basic Layer

### Transaction Patterns

- General rules
- Customized rules for specific scenarios
- Expert-defined rules and thresholds

### Account Profile

- Know your customer
- Watch list filtering
- Politically exposed individuals

## Graph Computing

- Flow analysis
- Advanced KYC, including the related parties of the customers
- Determines true relationships between accounts and parties
- Identifies important parties in networks and money flows

## Data Mining

- Pattern matching against previous cases
- Uncover links hidden in texts

## Grouping/Clustering

- By industry, business nature, region, size, etc.
- By transaction behavior
- By typologies, techniques

## Advanced Cognitive Layer

## Multi-Modality Analysis for Anomaly Detection

- Time-series analysis of transaction behavior and relationship change
- Behavior outlier detection against self history and peers
- Cognitive reasoning with Bayesian network inference

## Predictive Machine Learning

- Supervised learning to detect existing patterns
- Unsupervised learning to discover unknown patterns
- Assesses existing risk and predicts future risk



# **AI Regulation Reasoning & Compliance Reassurance**

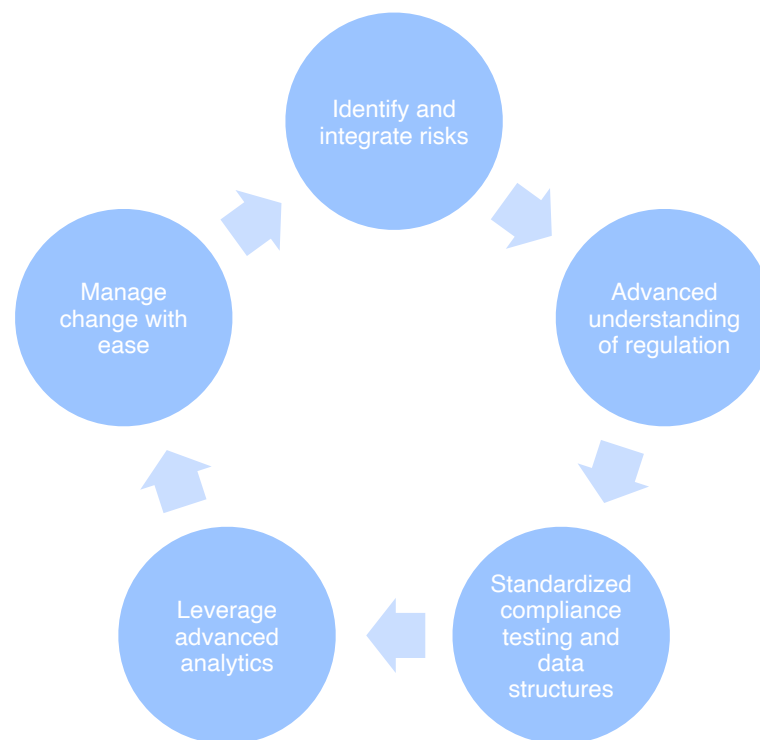
- The regulatory regime for Financial institutions are tightened worldwide. Banks are subjective to regulations such as CFPB, FRB, FinCen and others in U.S.
- Conventional compliance architecture is entangled in numerous systems, transformations, and mappings. At major banks each new compliance program brings **more** staffs, systems, software, warehouse and more documents.
- Cost fines peaked in 2016 at a total accumulated amount of over \$200 billion globally



# A snapshot of penalties for non-compliance issued recently

| Date       | News Title  | Penalties                            | Penalty type                  | Issued by         |
|------------|---|--------------------------------------|-------------------------------|-------------------|
| 03/07/2018 | <b>Bancorp Bank penalized \$2M for UDAP violations</b>              | \$ 2 million plus restitution        | UDAP/UDAAP                    | FDIC              |
| 02/15/2018 | <b>U.S. Bank NA paying \$598M for BSA/AML failings</b>              | \$ 598 million                       | BSA-AML Civil Money Penalties | FinCEN, DOJ, OCC  |
| 2/15/2018  | <b>U.S. Bancorp pays \$15M for BSA/AML failures</b>                 | \$ 15 million                        | BSA-AML Civil Money Penalties | FRB               |
| 2/7/2018   | <b>Rabobank pays \$369M for BSA/AML violations and obstruction</b>  | \$ 369 million                       | Forfeiture                    | OCC, DOJ          |
| 1/17/2018  | <b>Mega International Commercial Bank pays \$29M BSA penalty</b>    | \$ 29 million                        | BSA-AML Civil Money Penalties | FRB, State Agency |
| 12/27/2017 | <b>Citibank earns CMP for non-compliance with BSA C&amp;D order</b> | \$ 79 million                        | BSA-AML Civil Money Penalties | OCC               |
| 11/21/2017 | <b>Bureau fines Citibank for student loan servicing failures</b>    | \$ 2.75 million and consumer redress | UDAP/UDAAP                    | CFPB              |

- It is a subset of FinTech which utilize AI technologies such as Nature Language Processing(NLP), Nature Language Understanding (NLU) that may facilitate the delivery of regulatory requirements more efficient and effectively than existing capabilities. Regtech has emerged as the result of top-down institutional demand, in contrast to bottom-up demand that has driven FinTech.
- For banking industry, RegTech allows real time and proportionate regulation that identifies risk more efficient compliance and auditing procedures including:
  - Analyzing and implementing rules
  - Extracting, analyzing and storing data
  - Monitoring employee and customer behaviors





## Phase 1 - Manual

Manual data capture based on cyclical timelines (excel)

## Phase 2- Workflow automation

Compliance software establishes consistent workflow

## Phase 3- Continuous monitoring

Applying data science to automate the back office

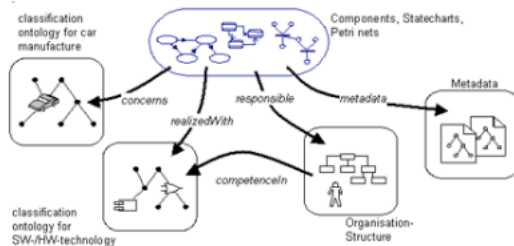
## Phase 4 - Predictive analytics

AI and machine learning are proactively identifying and predicting risk

# Master the Complexity of GRC (Governance Regulation Compliance) with Graph Approach



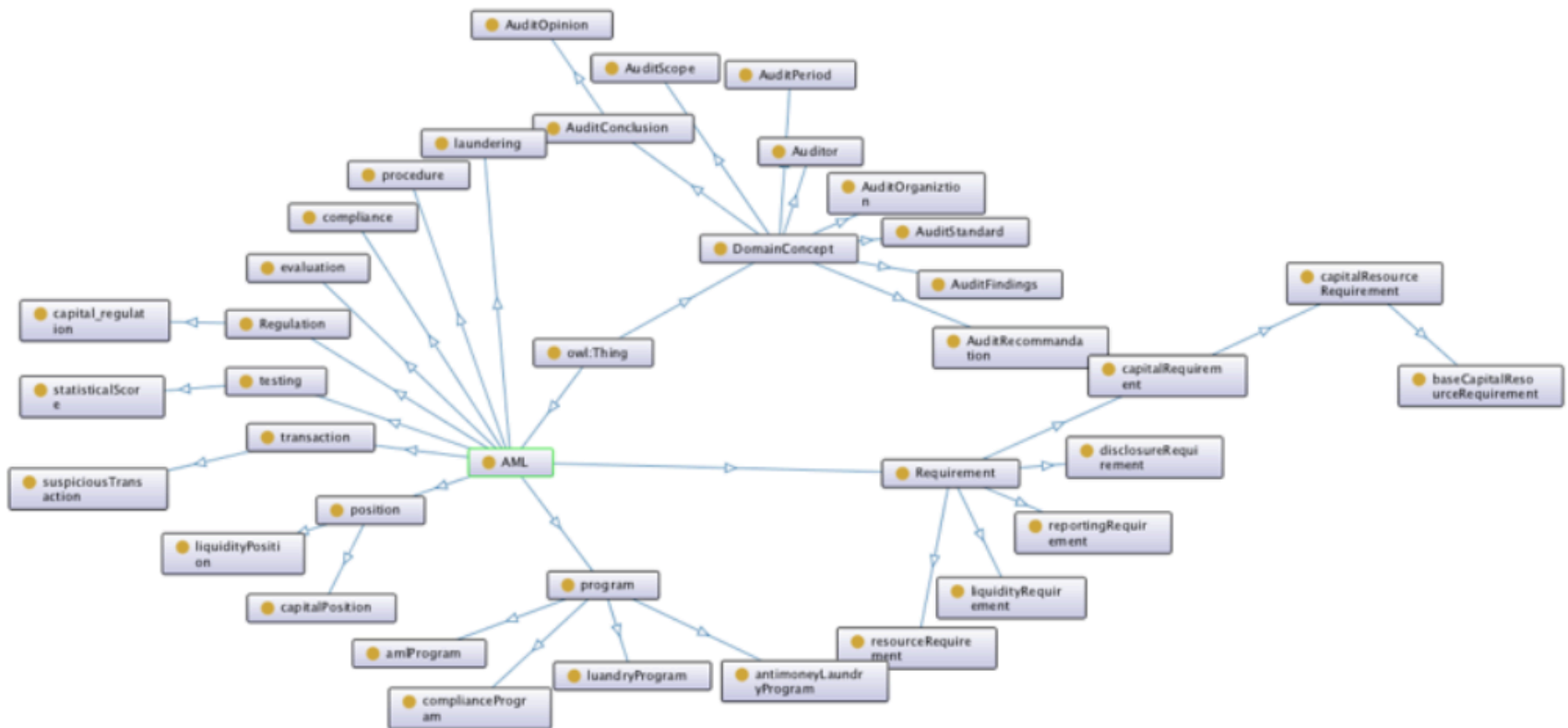
- Similar to the medical domain, finance can master the compliance challenge with graph
- The human genome contains more than 3 billion DNA base pairs. Genes direct the production of over a million analyzed proteins. More than 9500 terms define human phenotype and anomalies which describe over 10,000 disease. Almost half a million drugs are approved for treatment.
- The complexity of Bio/Medical is resolved with Semantic Web and Ontology, a graph representation of Triple System (subject/ predicate/ object) which facilitated understanding the relationships of terms.



Ontology graph from <https://www.w3.org/2004/12/rules-ws/paper/98/>

- AI technologies allow compliance professionals to “interpret regulatory meaning, comprehend what work needs to be done and codify compliance rules” in a fraction of the time normally required. AI can enhance compliance **monitoring, detection, and response**, incorporating forward-looking functions that **identify regulatory changes** and enable businesses to update procedures quickly.
- Extracting metadata: NLP identifies important elements of a regulation and helps users to understand what the document is about.
  - If the regulation is relevant
  - How the organization may be affected and needs to respond
- Identifying entities: NLP can determine the “who” factors in regulation:
  - To whom the document is addressed (such as a firm or department)
  - By whom (such as a regulator)
  - Who are the key actors (such as customers or market participants)
- “Understanding” content: NLP can help users to
  - identify the requirements that are contained within a document
  - using the entities and metadata, determine who they apply to and what products, topics and processes they refer to

# Example: AML regulation analysis for auditing



(CNN)French striker Bafetimbi Gomis, who has a history of fainting, said he is now "feeling well" after collapsing during Swansea's 3-2 loss at Tottenham in the Premier League on Wednesday. The worrying incident occurred in the first half at White Hart Lane -- after Tottenham scored in the seventh minute -- but the 29-year-old left the pitch conscious following about five minutes of treatment. The Guardian added that he was wearing an oxygen mask. Play was temporarily stopped before resuming. Fabrice Muamba on playing soccer again 00:56 As the match progressed, Swansea tweeted that Gomis was "fine," with manager Garry Monk using the same word to describe Gomis' condition. Gomis spent the night in hospital as a precaution, Swansea said on its website.

I wanted to reassure you concerning my health," Gomis told the website. "It actually looks much scarier than it is physically dangerous, and I am feeling well now. "I have been under a great deal of stress and fatigue due to my father's health, which requires me to go back and forth from France. "I was disappointed that I couldn't help my team tonight, but now everything is back in order. I also want to thank everyone for their support and get well messages."

Gomis had similar fainting spells in France, which prompted the president of his former club, Jean-Michel Aulas of Lyon, to tell French television in 2009: "We can't not be worried, it scares you each time."

Swansea ran tests on Gomis, said Monk, prior to signing him on a free transfer last July.

"He just has a little bit of low blood pressure which causes you a little bit of problems," Monk said in a televised interview on Sky.

"It's been part of his life. We were well aware of that when we signed him. He's done all the hospital checks and all the medical checks you can possibly do and it's just part of his life.

"It's no problems whatsoever. It's not as serious as it looks."

Gomis has scored two league goals for Swansea this season, mostly in a backup role. He became the Welsh side's top striker when Wilfried Bony signed with Manchester City in January.

Almost exactly three years ago at White Hart Lane, then Bolton midfielder Fabrice Muamba collapsed after suffering a cardiac arrest. He was near death, according to Bolton, but survived after being treated at the London Chest Hospital.

He subsequently retired.

Other footballers, including Cameroon international Marc-Vivien Foe in 2003 and Spanish international Antonio Puerta in 2007, didn't survive after collapsing on the pitch.

News from CNN

## Summarization Results:

Reference Summary: Bafetimbi Gomis collapses within 10 minutes of kickoff at tottenham, but he reportedly left the pitch conscious and wearing an oxygen mask. Gomis later said that he was "feeling well" the incident came three years after Fabrice Muamba collapsed at white hart lane .

## Generated Summary:

Bafetimbi Gomis says he is now " feeling well" after collapsing during Swansea 's 3-2 loss . the 29-year-old left the pitch conscious following about five minutes of treatment . The 29-year-old left the pitch conscious following about five minutes of treatment .



'Summary of Business Nature of Business Energy West was originally incorporated in Montana in 1909 and was reorganized as a holding company in 2009. On July 9, 2010, we changed our name to Gas Natural Inc. (the "Company," "we," "us," or "our") and reincorporated in Ohio. We are a natural gas company with operations in four states. In October 2016, we implemented a plan of reorganization and formed a new holding company, PHC, an Ohio Corporation, that is the parent company of our regulated utility subsidiaries, Cut Bank Gas, EWM, Frontier Natural Gas, Bangor Gas, NEO, Brainard, Orwell, and Spelman. Gas Natural is the parent company of Energy West Propane, Inc., EWR, GNR, Lone Wolfe and PHC. PHC is the parent company of multiple entities that are natural gas utility companies with regulated operations in Maine, Montana, North Carolina and Ohio. EWR is a natural gas marketing and production company with non-regulated operations in Montana. GNR is a natural gas marketing company that markets gas in Ohio. Energy West Propane, Inc. distributes propane with non-regulated operations in Montana. Lone Wolfe serves as an insurance agent for us. We have three operating and reporting segments: • Natural Gas. Representing the majority of our revenue, we annually distribute approximately 21 Bcf of natural gas through regulated utilities operating in Maine, Montana, North Carolina and Ohio. Our natural gas utility subsidiaries include Bangor Gas (Maine), Brainard (Ohio), Cut Bank Gas (Montana), EWM (Montana), Frontier Natural Gas (North Carolina), NEO (Ohio) and Orwell (Ohio). As of December 31, 2016, we served approximately 69,400 customers. • Marketing and Production. Annually, we market approximately 3.6 Bcf of natural gas to commercial and industrial customers in Montana, Wyoming and Ohio through our EWR and GNR subsidiaries. Our EWR subsidiary also manages midstream supply and production assets for transportation customers and utilities. EWR owns an average 53 % gross working interest (average 44 % net revenue interest) in 160 natural gas producing wells and gas gathering assets located in Glacier and Toole Counties in Montana. • Corporate and Other. Included in corporate and other are costs associated with business development and acquisitions, dividend income, recognized gains or losses from the sale of marketable securities, activity from Lone Wolfe which serves as an insurance agent for us and other businesses in the energy industry, and the results of our discontinued operations from the sales of EWM, Pipeline Assets and Independence. '

## Topic:

Business\_Finance  
Environment

## Industry:

Auto, Truck & Motorcycle Parts  
Oil & Gas Exploration and Production  
Natural Gas Distribution  
Food Markets  
Wires & Cables

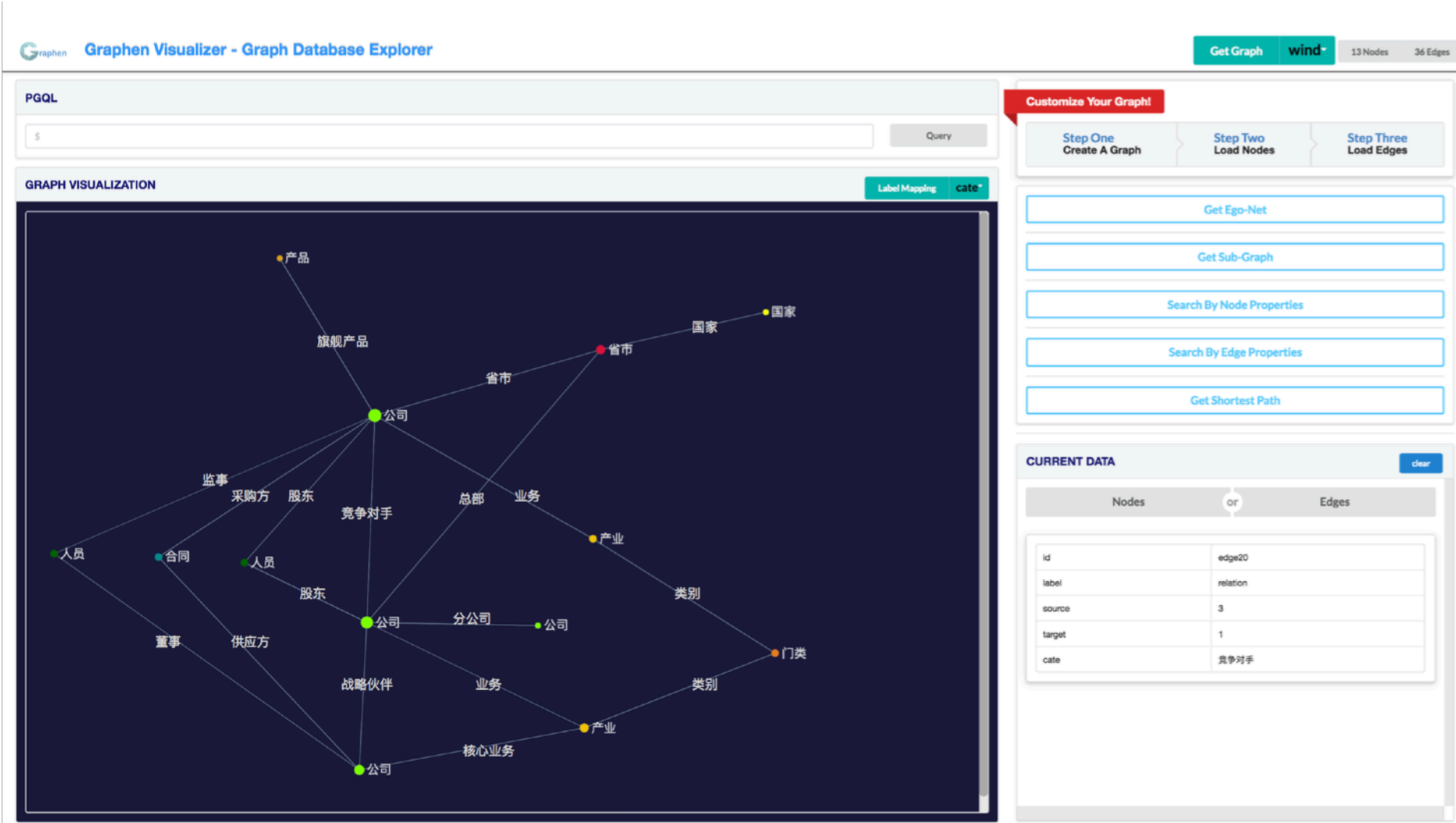
## Relations:

Company Affiliates  
Company Name Change  
Company Reorganization  
Personal Career  
Company Founded

## Tags:

Lone Wolfe  
Natural gas  
Economy of United States  
Energy  
Economy of North America

- Graph analysis by Graphen Ardi platform



# A Comprehensive GRC(Governance Regulation Compliance) Knowledge Graph

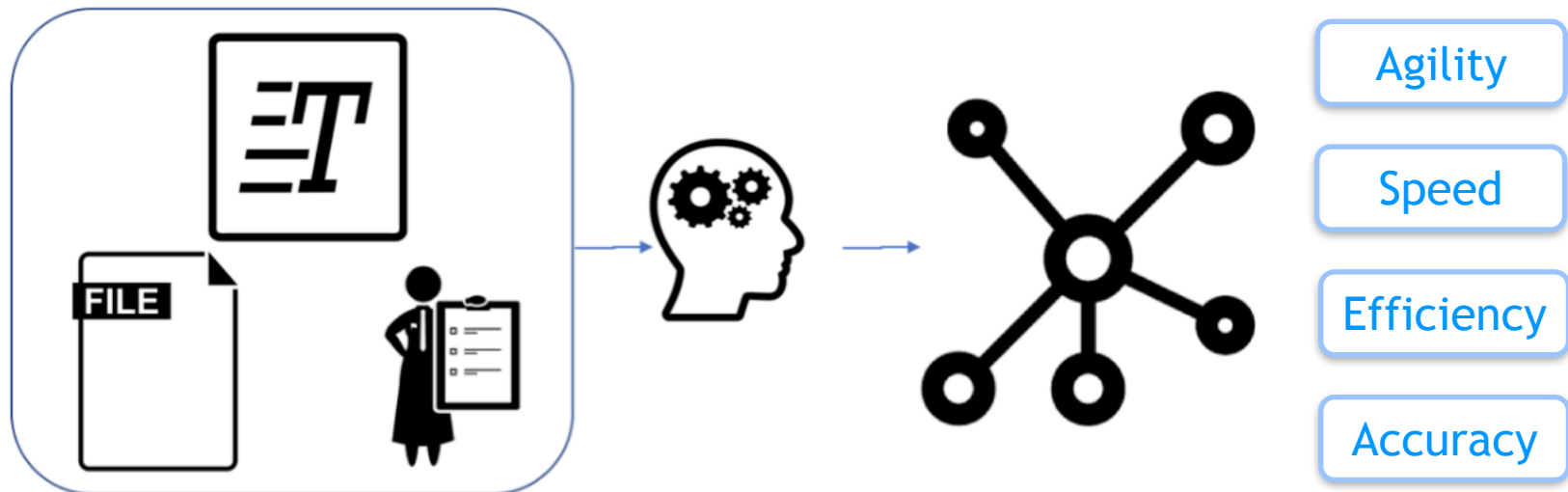
Challenge: Compliance responsibilities are spread throughout the organization so that risk assessment, testing and reporting lacks integrated data with same formats which makes internal compliance and audit process efforts slow and expensive

Solution:

**Integrated Data:** Design algorithms to extract and integrate data from banks' proprietary system, third-party data providers, regulations, regulatory announcement and public sources including both structured and unstructured data.

**Cognitive learning:** Utilize both machine learning and human expertise to continuously improve the quality, precision and reliability of data.

**Graph Representation** Refine the design the domain and linkage of data, and store the data in the powerful graph.



## Improve Regulation Compliance

Understanding the Regulation is the key for three lines of defenses

Graphical representation of regulation as the standard graph

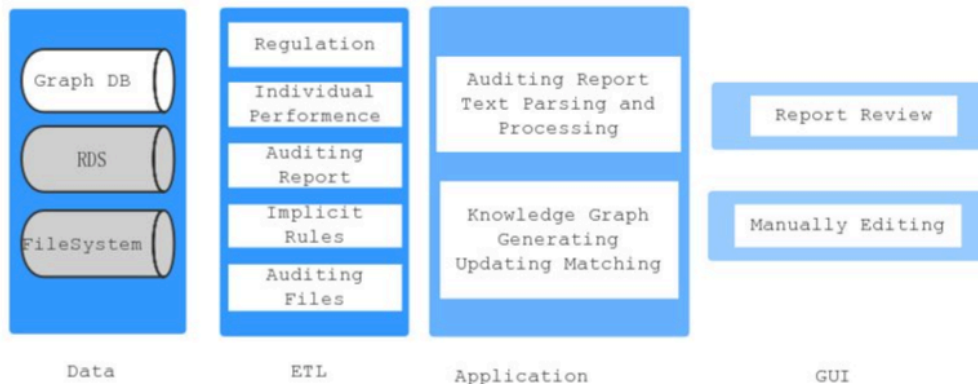
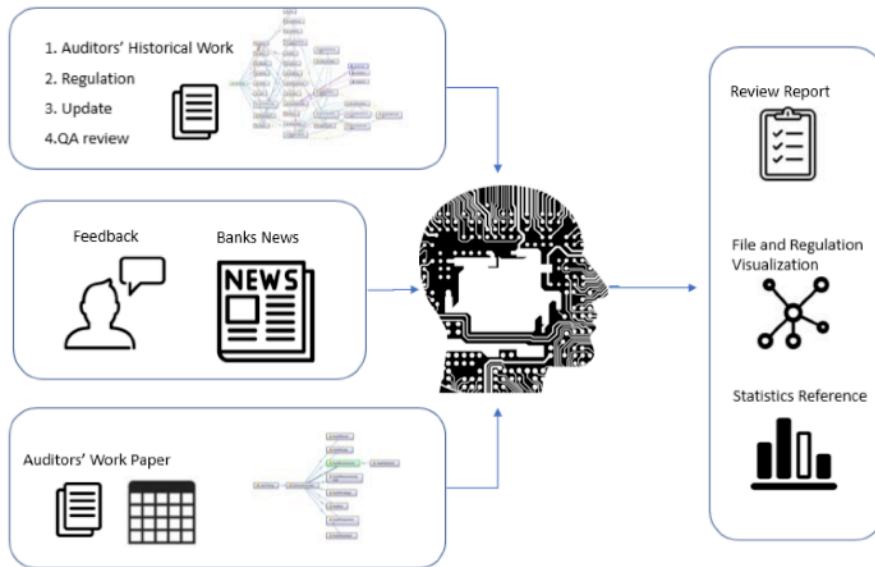
Extract information from internal data such as policy controls, procedures and supporting documents corresponding to regulation terms and represent in the internal-data graph

Compare the standard graph with internal-data graph and check the missing requirement and possible violations

Linked data from different domains increase efficiency and reduce cost of repetitive works

## Key Functions of Graphen RegTech

- Deep understanding of regulation and your organization
  - Key issues such as KYC, AML, Customer protection services
  - Regulatory agencies
  - Penalty amounts
  - Regulatory requirements (documents, procedures and controls)
  - Related departments
  - Enforcement actions
- Monitoring, detecting and response to compliance risks
  - System will find the missing requirements for given issue
  - Identify other risks regarding the similar problematic internal data
  - Give recommendation for risk reporting, and enable feedback to improve the system

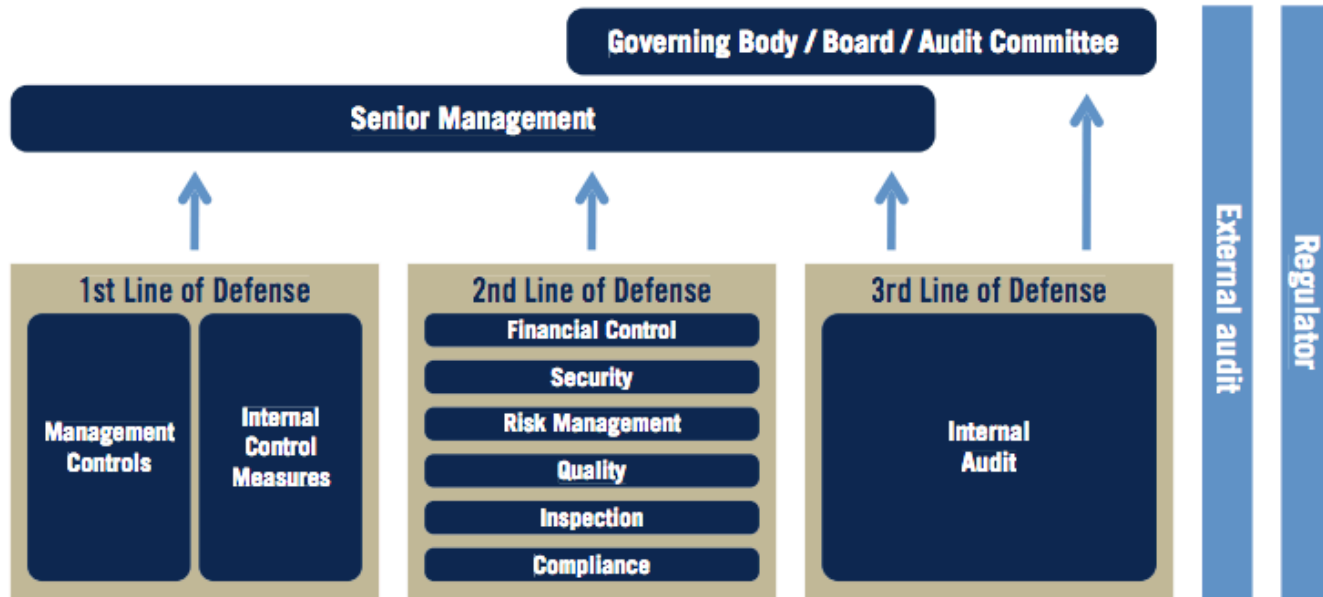


- Key Features:
  - Knowledge Graph based advance regulation comprehension
  - Semantic search of related regulation
  - Auditing report ELT and data mining
  - Report quality assessment
- Financial data, regulations, reports, workpapers and metadata can be stored in a uniform way. A knowledge graph defines the semantics of concepts, their relationships and axioms. Compliance crosses the domains of finance and legal regulations.
- To improve the quality of the auditor's report, the system give recommendation based on
  - Graph comparison of regulation and internal data
  - Other auditor's behaviors analysis as benchmark
- The system keep learning by user's feedback



- Traditional Approach

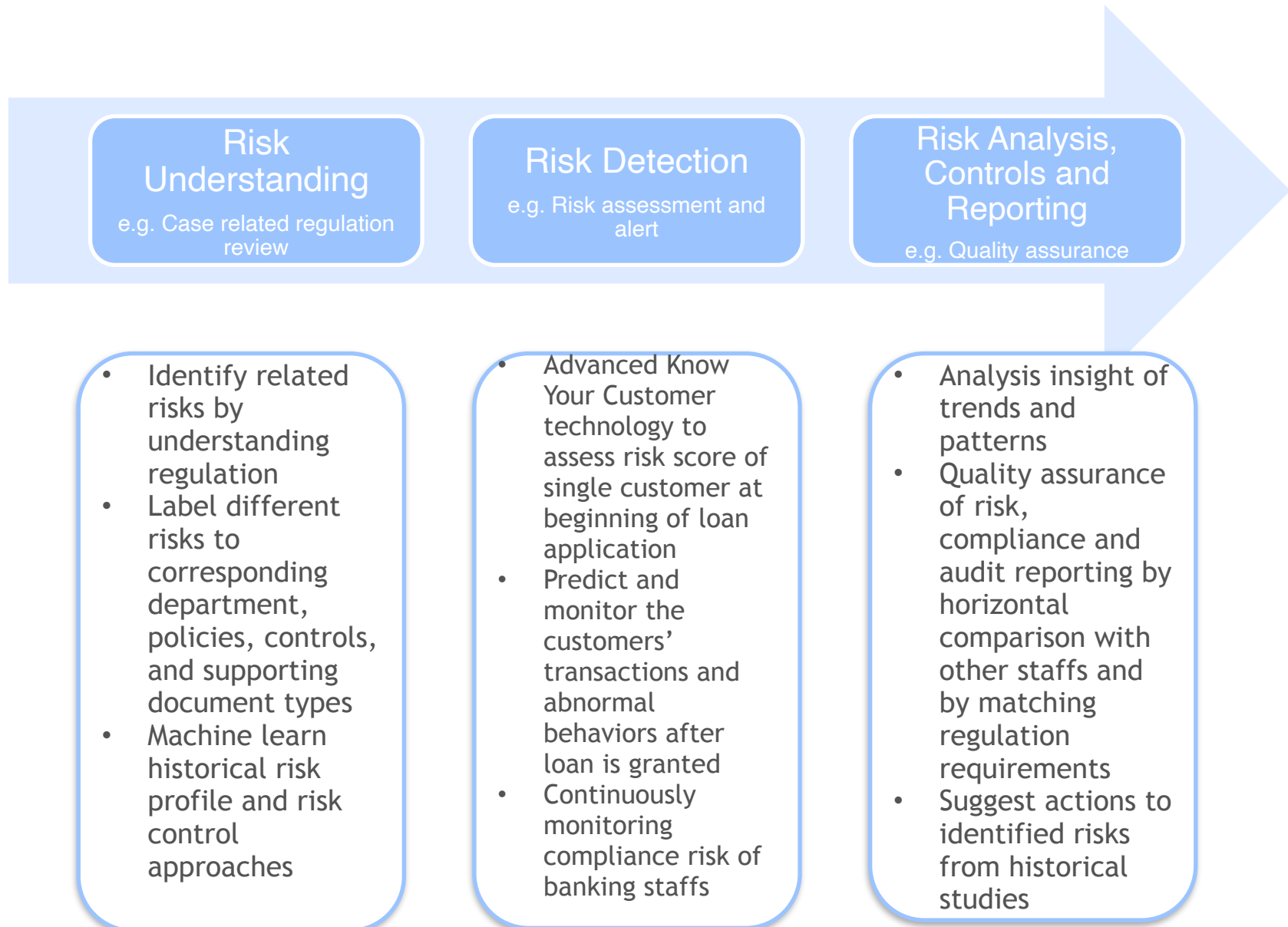
## The Three Lines of Defense Model



Graph Adapted from ECIIA/FERMA Guidance on the 8<sup>th</sup> EU Company Law Directive, article 41

- Graphen AI risk defense in one-stop







**Market Intelligence**

- **What is Market Intelligence?**
- **Market Intelligence Platform Functions**
- **How Market Intelligence Platform Analyzes A Single News?**
  - **Bayesian Network**
  - **The Science Behind Reasoning**
- **Price Impact Prediction**
  - **How Market Intelligence Platform Aggregates Information?**
  - **Aggregate Information Analysis Framework**

- **Personal research analyst**
- **Provides real-time news information with relevance rankings according to your individual portfolios and watchlists**
- **Generates overall Price Impact Score from all monitored news sources**



- Total impact: calculated from all monitored news source
- Single News Impact: calculated from selected news

Artificial intelligence and Bayesian network powered causality reasoning graph



What's in  
your  
portfolio

Price Chart

Trending  
news that  
has impact  
on your  
portfolio

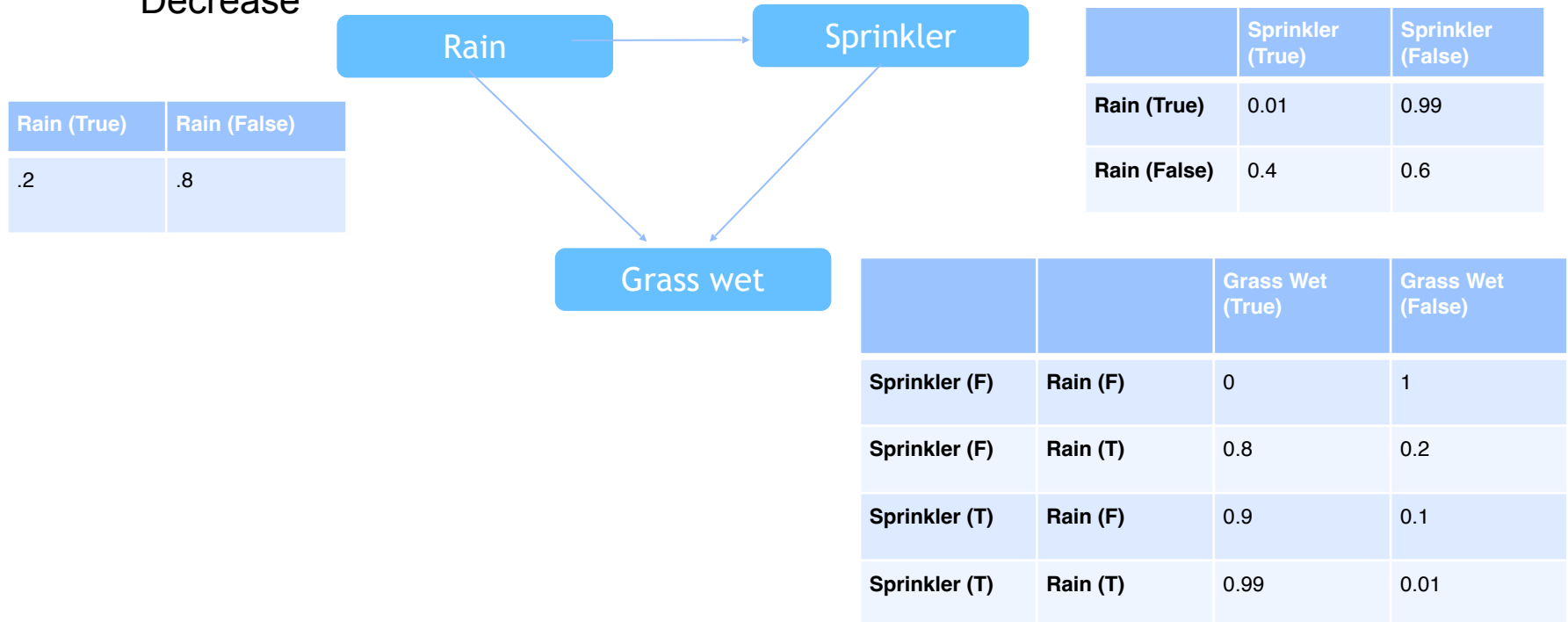
Personalization of  
news you care

Bayesian network is a probabilistic graphical model (a type of statistical model) that represents a set of variables and their [conditional dependencies](#) via a [directed acyclic graph](#) (DAG)

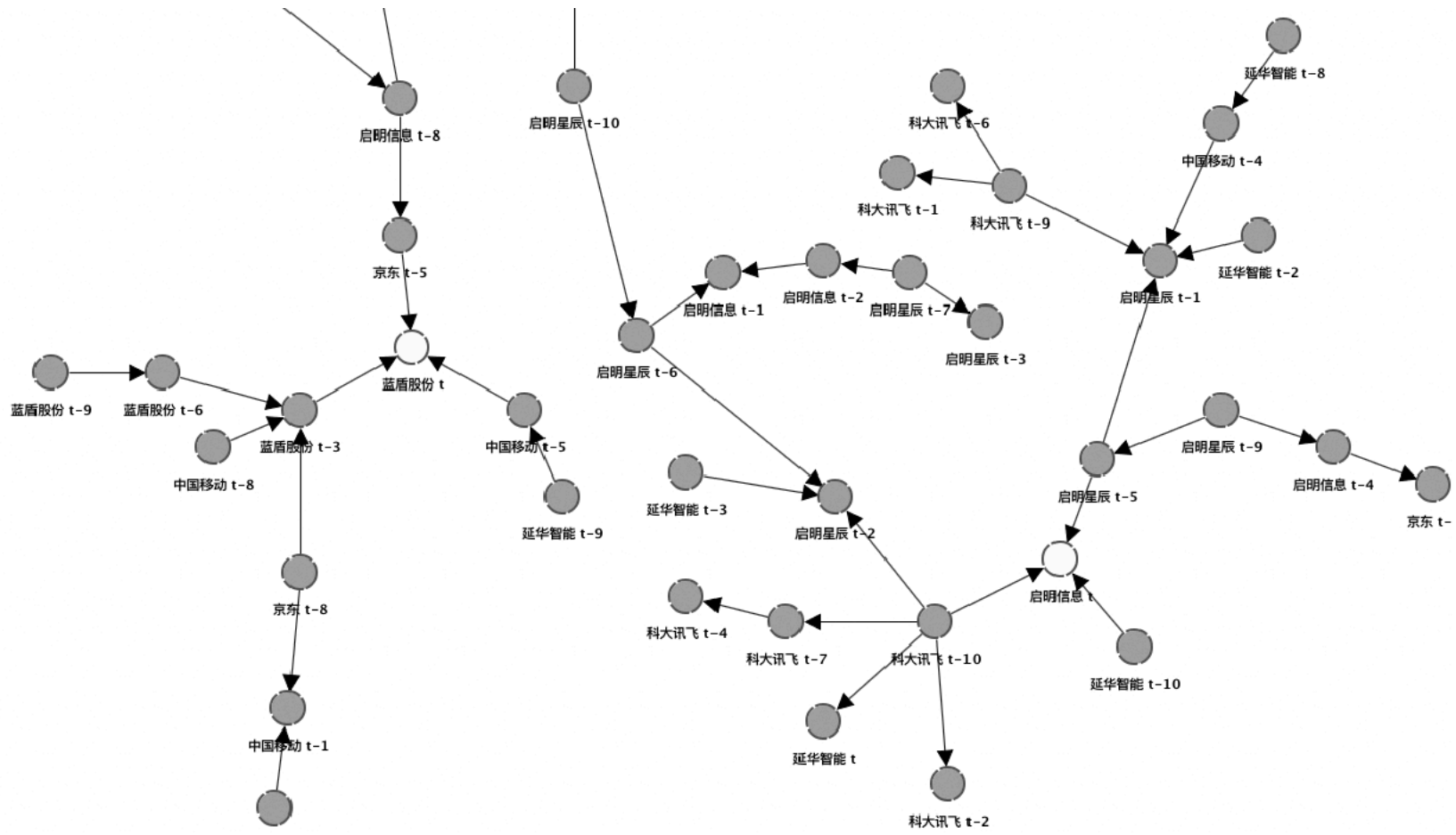
A simple Bayesian network. Rain influences whether the sprinkler is activated, and both rain and the sprinkler influence whether the grass is wet.

With observed probability, we can answer questions such as what is the probability of raining, given grass is wet.

Similar causality network can be applied to the market: Rain – Negative Economy Outlook; Sprinkler – Negative Company Management ; Grass Wet – Stock Price Decrease



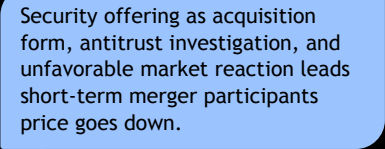
| Ticker | Company | Industry                  | Sector |
|--------|---------|---------------------------|--------|
| 002178 | 延华智能    | Information Technology    | Tech   |
| 002230 | 科大讯飞    | Information Technology    | Tech   |
| 002232 | 启明信息    | Information Technology    | Tech   |
| 002439 | 启明星辰    | Internet Service Provider | Tech   |
| 300297 | 蓝盾股份    | Information Technology    | Tech   |
| JD     | 京东      | Internet Service Provider | Tech   |
| CHL    | 中国移动    | Information Technology    | Tech   |



# How Market Intelligence Platform Analyze A Single News?

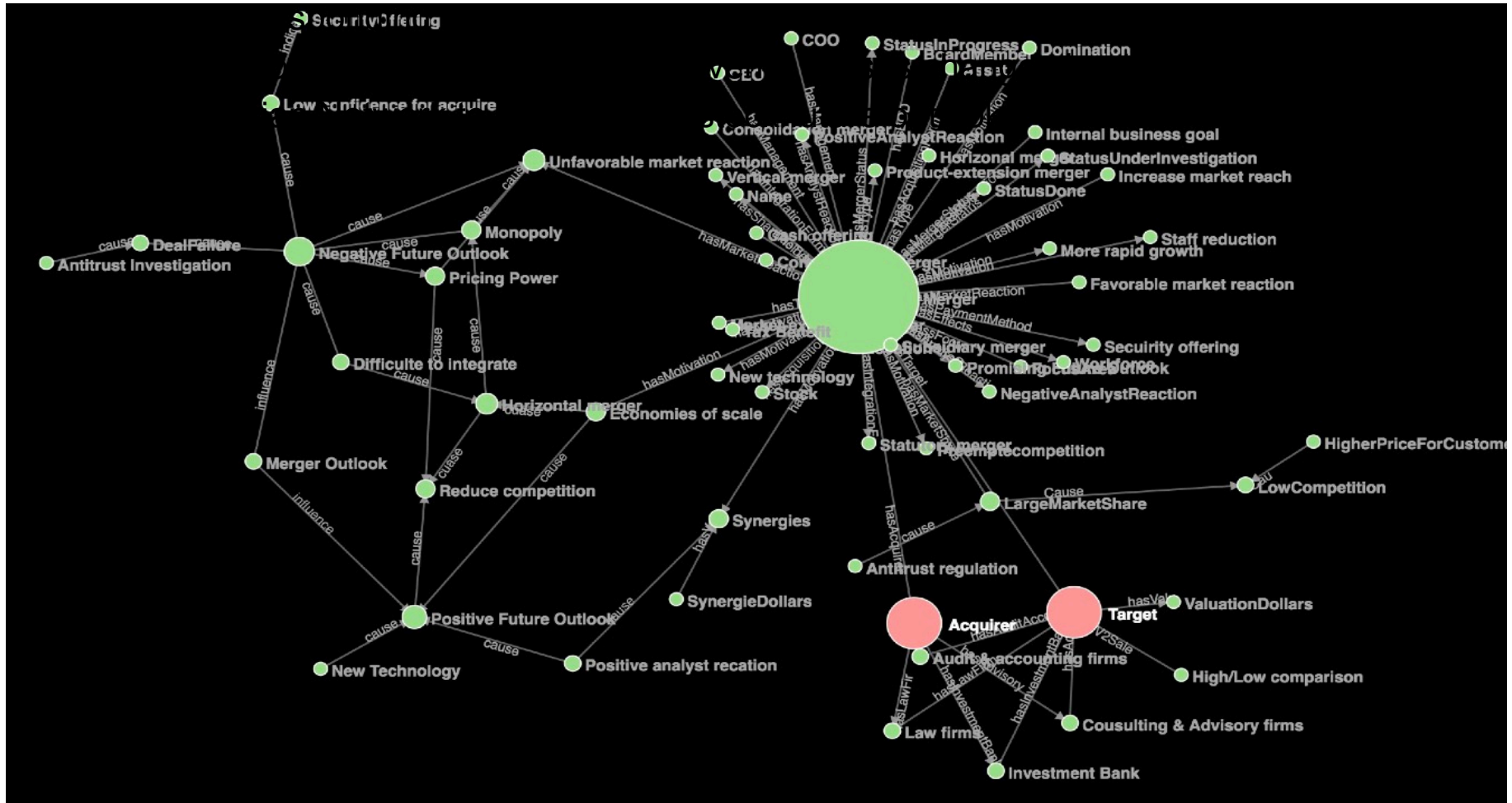
Impact score of  
news on each  
stock

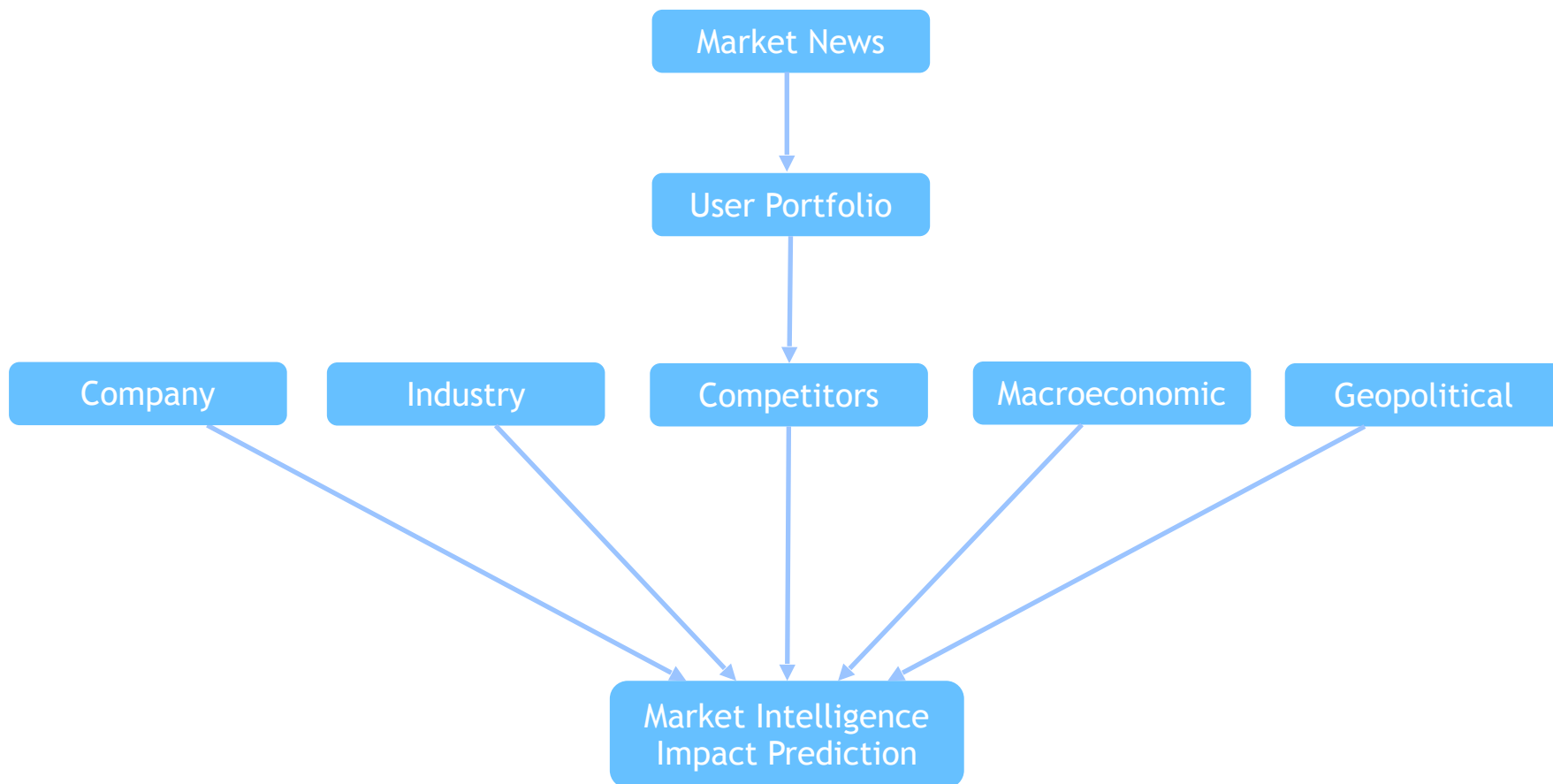






## Event analysis module for each type of events





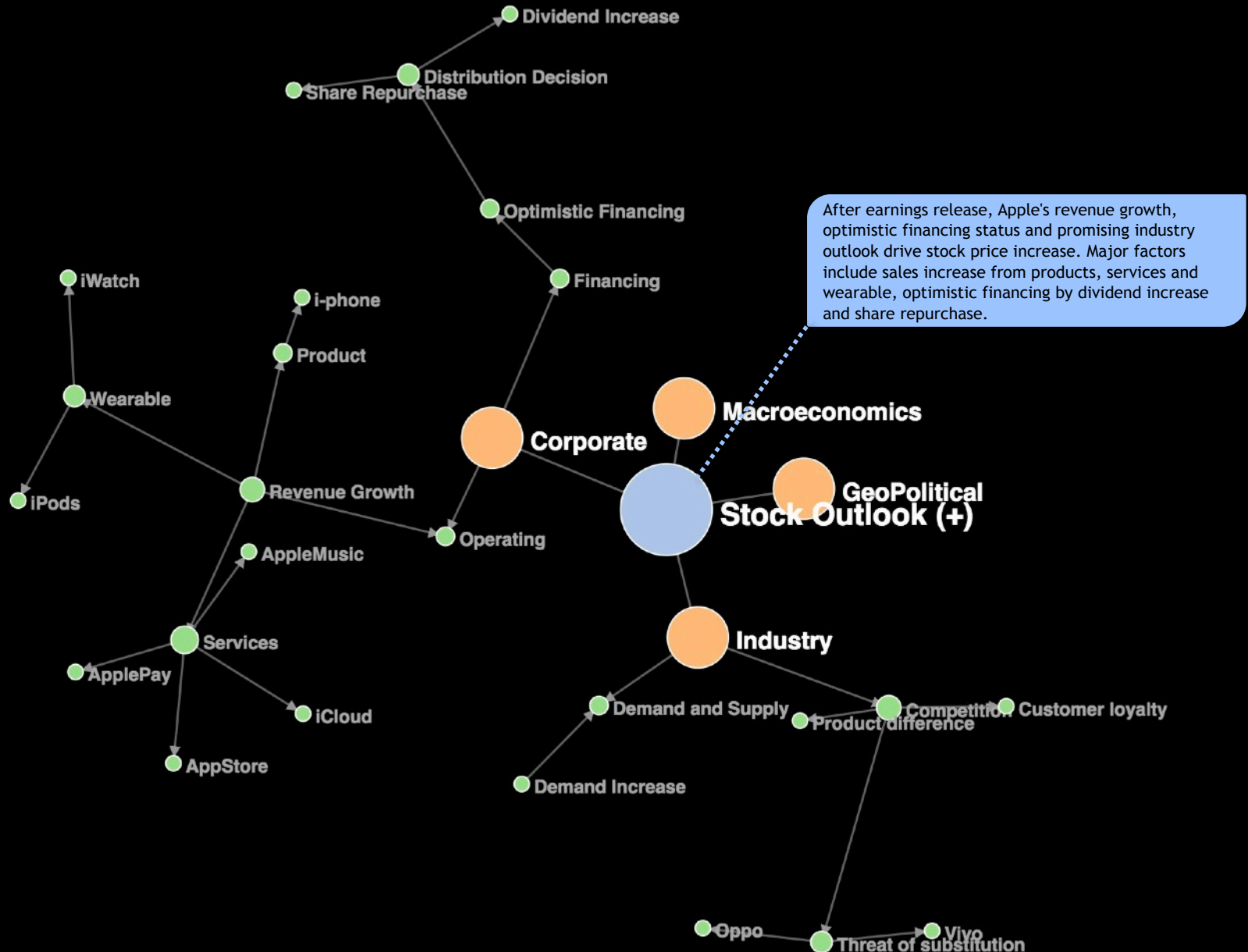
# How Market Intelligence Platform Aggregate Information?

Assess  
aggregate  
information of  
Apple Inc. stock



- Stock price predicted to be increase by aggregate all monitored news from four major areas.
- Detailed news reasoning graph of recommended news from left column.

# Detailed Apple Stock Reasoning Graph



- Corporate

- Investing: external investing, internal investing
- Operating: revenue, cost, management, product
- Financing: capital structure, distribution, debt

- Industry

- Lifecycle
- Demand Supply
- Future expectation

- Competition

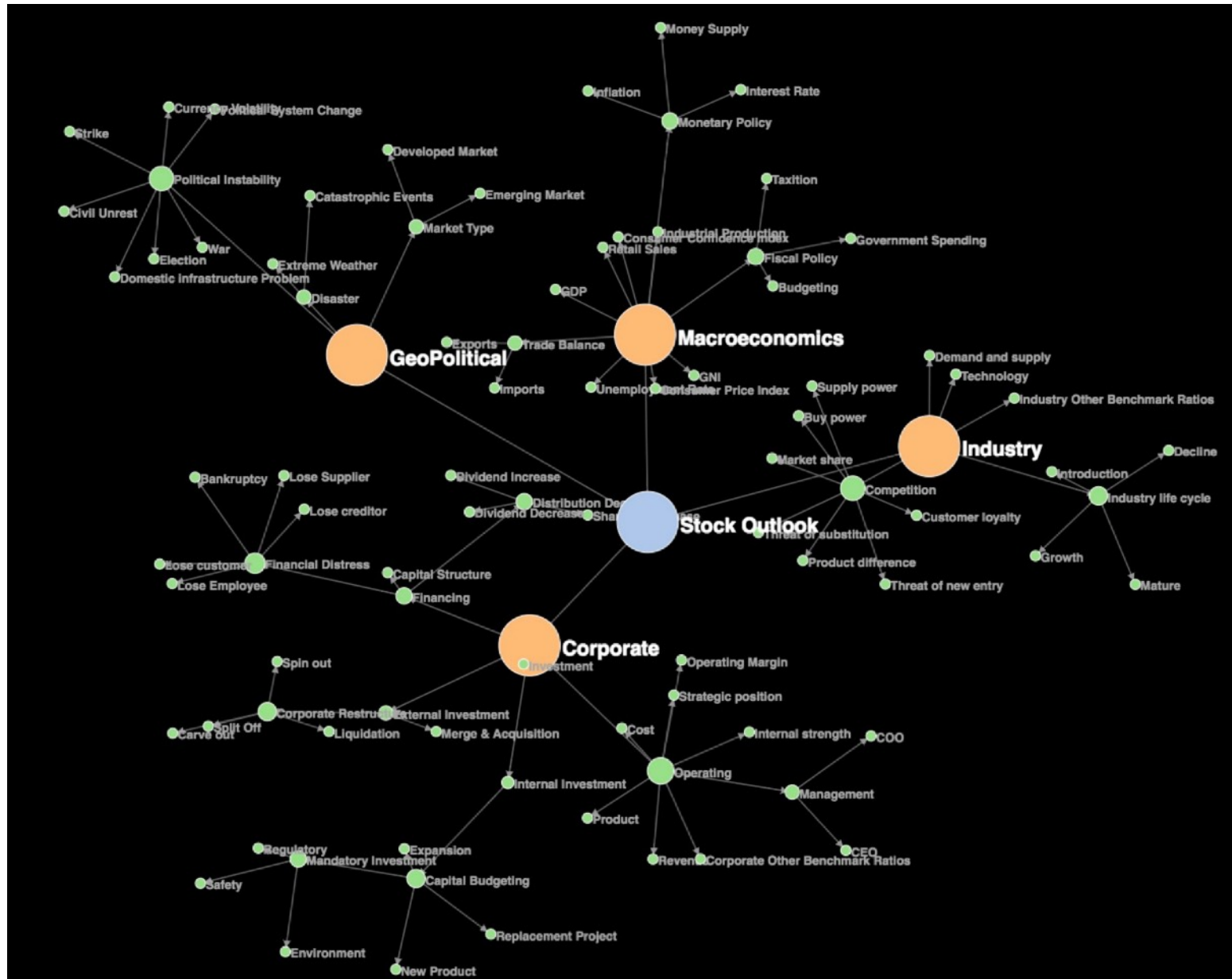
- Market share
- Product differentiation

- Macroeconomic

- Indicators such as GDP, GNI, Retail Sales, Unemployment Rates, CCI and etc.
- Trade Balance
- Monetary Policy: Inflation, Interest Rate, and etc.
- Fiscal Policy: Taxation, Government budgeting

- Geopolitical

- Market type: Emerging, Developed
- Nature disaster, extreme weather, catastrophic event
- Political instability: war, strike, civil unrest








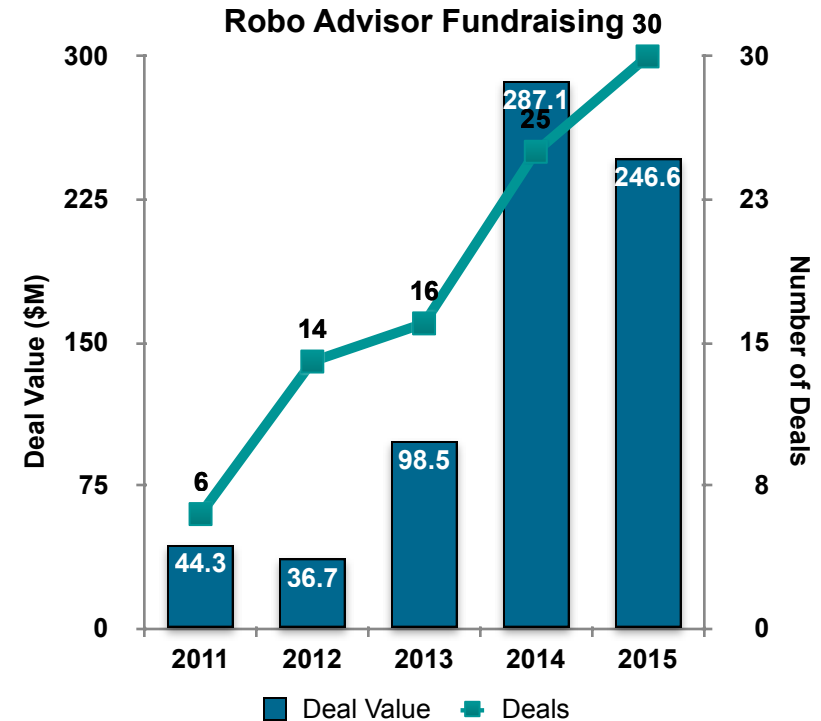




# AI for Robo-Advisory

# Robo-Advisory Examples

| Type                   | Description  | Example   | Others  |
|------------------------|--|---|---|
| Robo Advisors advisors | Digitalized automatic wealth management platform   |    | 4Q2015: Top 5 Robo Advisors have combined AUM being \$44.2B.  |
| Retail Investment      | New type of digital transaction and investment platform, such as theme investment or combined transactions |    | Robinhood boasts a sign-up time of 4 minutes and free trading |



Sources: Investmentnews.com,

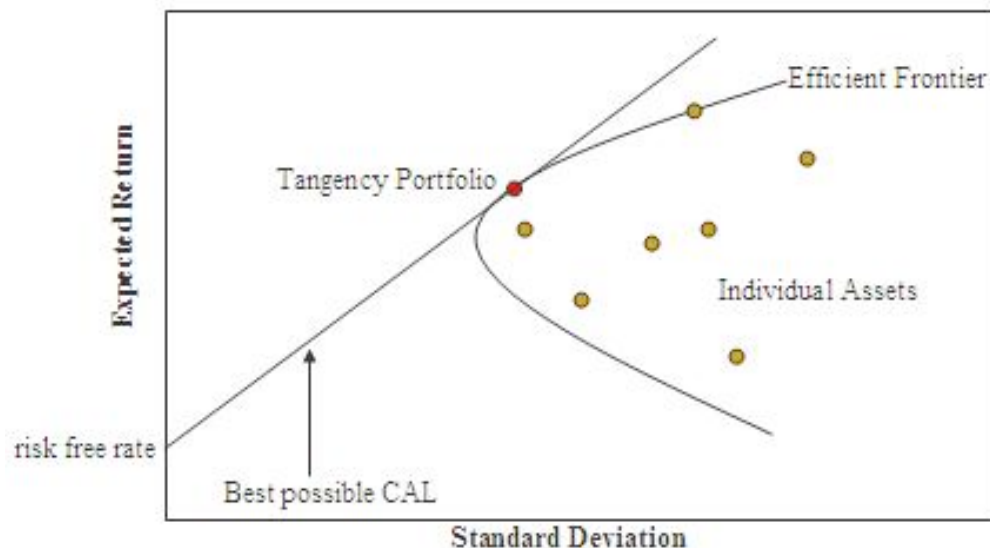
Robo-Advisor is a new type of wealth management service. Based on the risk level and investment goals provided by the investor, and it uses a series of 'smart algorithm' to calculate the optimal investment suggestions.

Robo-advisors directly managed about \$19 billion as of December 2014. By 2020 the global assets under management of robo-advisers is forecast to grow to an estimated US\$255B.

- Non-biased
- Low investment threshold
- Low starting entry money
- Low agent fee

## Features :

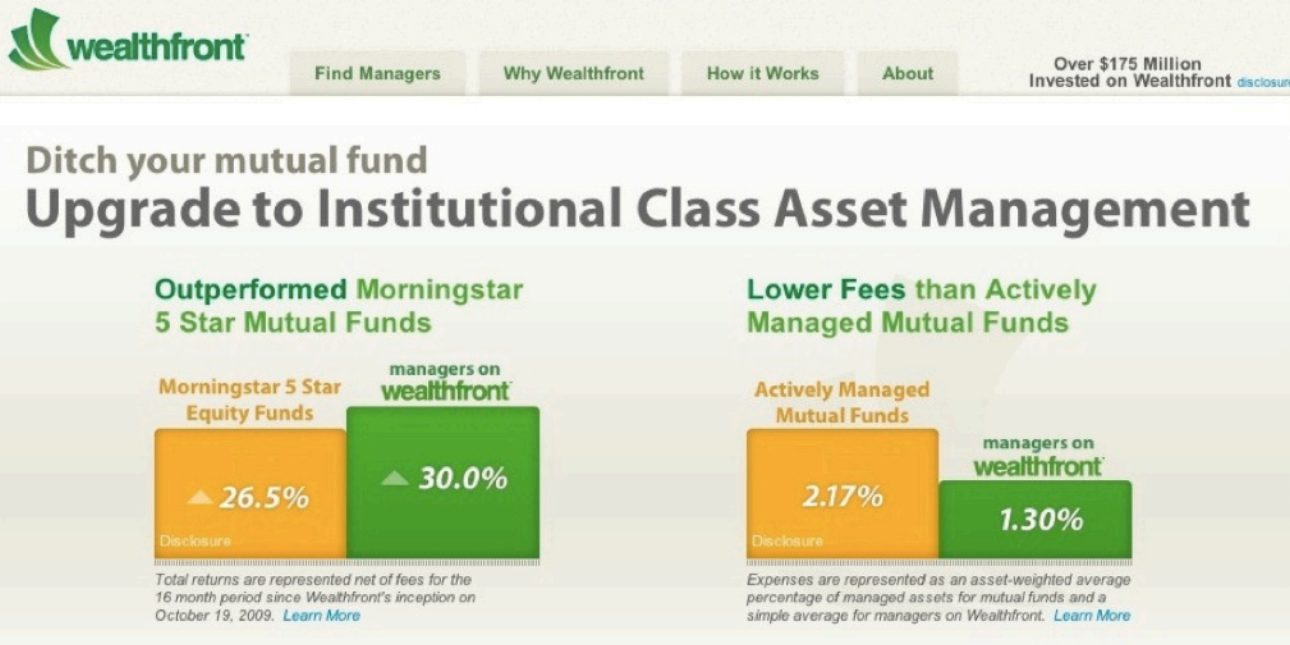
- Strongly depend on technology, algorithm and financial theory
- Distributed investment, maximum long-term return
- Personalized portfolio allocation.



Harry Markowitz

## Example: Wealthfront—low entry requirement, low fee

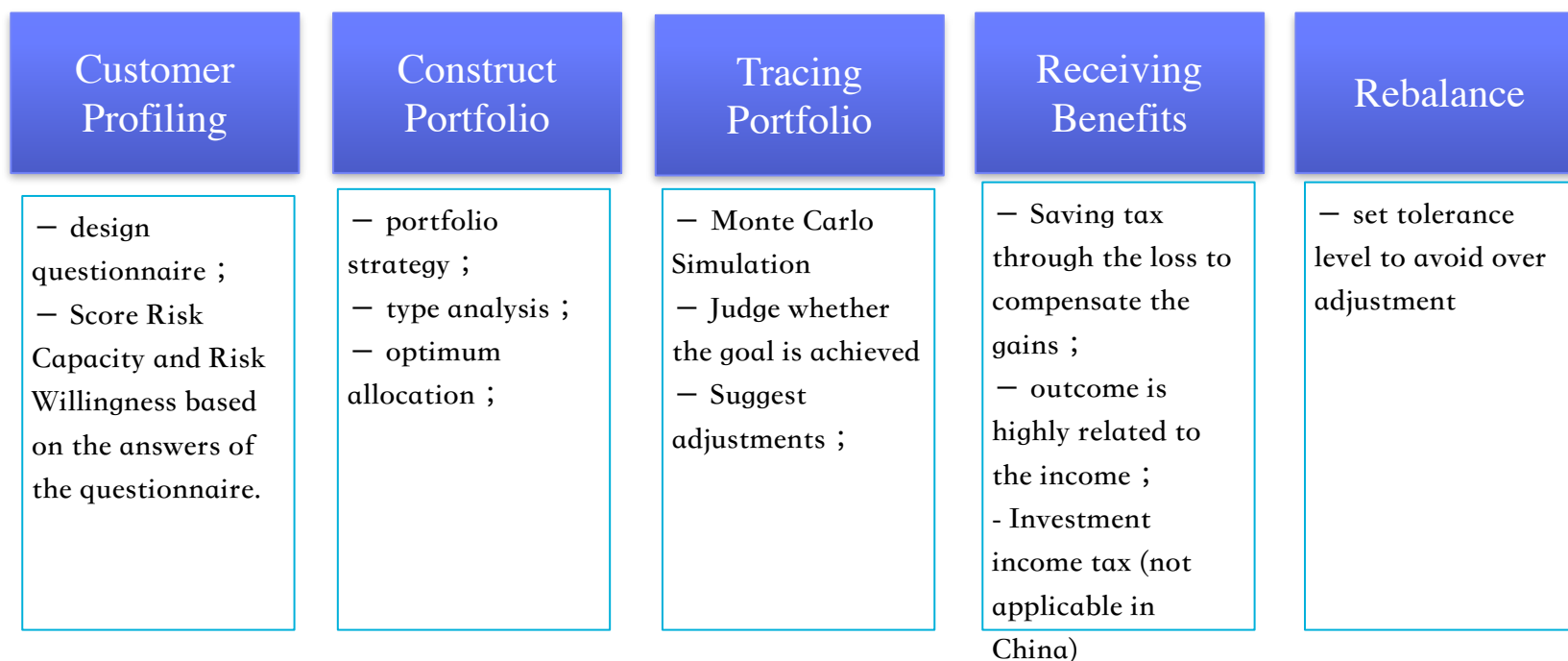
- Established in 2008. Formally called 'Kaching'. Renamed in Dec 2011. Headquarter at Palo Alto.
- On Sept 2015, the total asset is \$2.6B.
- The estimated value of the company is \$1B as in 2015.



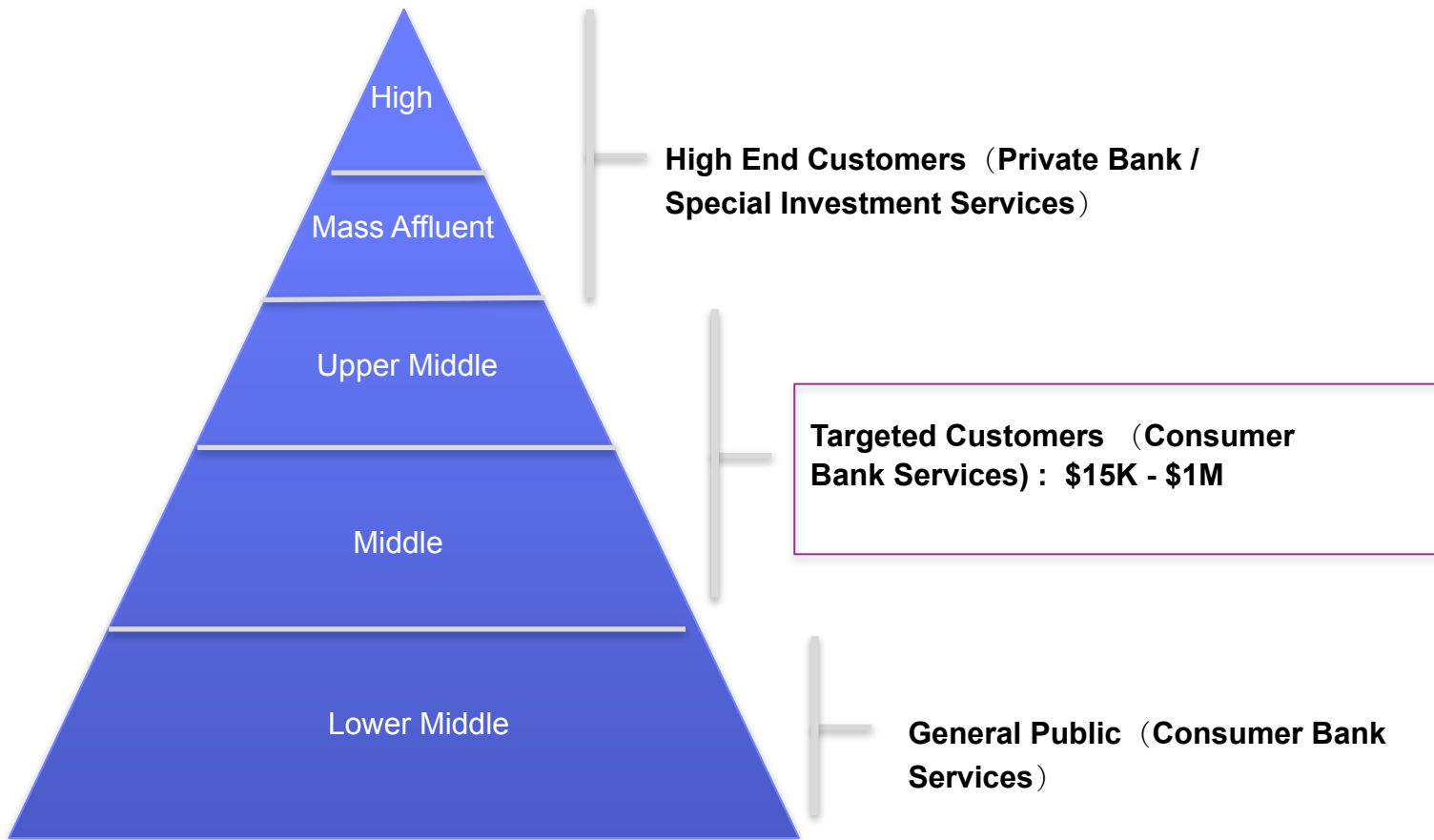
- **Low entry requirement** : min investment value USD \$500.
- **Low fee** :
  - Zero annual fee, if account is lower than \$10K.
  - 0.25% fee is charged for the part of asset amount that is larger than \$10K.
  - No agent fee
  - Based on Wealthfront, in average, each user only needs to pay 0.12% fee.



Most of the robo-advisor platform is built based on the modern investment portfolio theory, using Exchange Trade Funds (ETFs) to build portfolio.



Based on a survey of Wells Fargo, in US, there is only 16% of population in their 20s and 30s are willing to interact with investment consultants. The remaining people prefer to use these types of AI consultant.





## Four Steps to use Big Data Cognitive Analysis for Robo-Advisor

| Investment Market Analysis  | Dynamically Know Your Customer   | Optimized Personalized Investment Strategy   | Precise Bank-Customer Interaction   |
|---|--|--|---|
| <ul style="list-style-type: none"> <li>Analyze the market performance of various kinds of funds</li> <li>Analyze domestic and international financial and economic changes and how they may impact CPI, PPI, or GDP.</li> <li>Use Machine Learning and Deep Learning, based on historical economic numbers, find out how factors impact financial markets.</li> </ul> <p><u>Data</u></p> <ul style="list-style-type: none"> <li>Product Data</li> <li>Market Data</li> <li>Historical Economic Data</li> <li>Industry-related Data</li> </ul> | <ul style="list-style-type: none"> <li>Customer Profiling, e.g, based on IPQ (Individual Profile Questionnaire), Feedback, Risk Capacity and Risk Willingness</li> <li>Understand what the customer really wants based on their past behaviors interacting with bank</li> </ul> <p><u>Data</u></p> <ul style="list-style-type: none"> <li>Customer Data</li> <li>Behavior Data / Interaction Data</li> </ul> | <ul style="list-style-type: none"> <li>Strategy computation and optimization based on personal history</li> <li>Demonstrate / Simulate 'what ifs' when the portfolio has different allocation.</li> <li>Explainability of 'what ifs' to customer to the customer.</li> </ul> <p><u>Data</u></p> <ul style="list-style-type: none"> <li>Customer Data</li> <li>Market Data</li> </ul> | <ul style="list-style-type: none"> <li>Create and predict customer interaction strategy, including when, method, content to interact with customer – to achieve max customer and bank benefit.</li> </ul> <p><u>Data</u></p> <ul style="list-style-type: none"> <li>Customer Data</li> <li>Interaction Data</li> <li>Else?</li> </ul> |

# Example: Major Wealth Management



**Hundreds of  
products/campaigns  
Combinations with  
incompatibilities**

**How much of each product/  
campaign ?**



**Nightly batch run,  
select over 1.2M**



**Experts doing what-if  
to improve process**



**Telesales,  
Mail, email,  
Office, etc...**

**Done through which  
channel?**



**To which customers?**

**Several millions of  
customers**

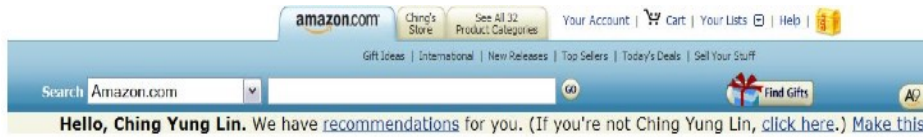
**When?**

**Select actions for the  
next days**



- Using customer past investment transactions under different market conditions, determine customer risk taking sentiment in real time on daily basis. Is customer panicking type or double down?
- Assess customer financial strength in taking aggressive investment position.
- Suggest portfolio adjustment at a rate that matches customer investment change rate

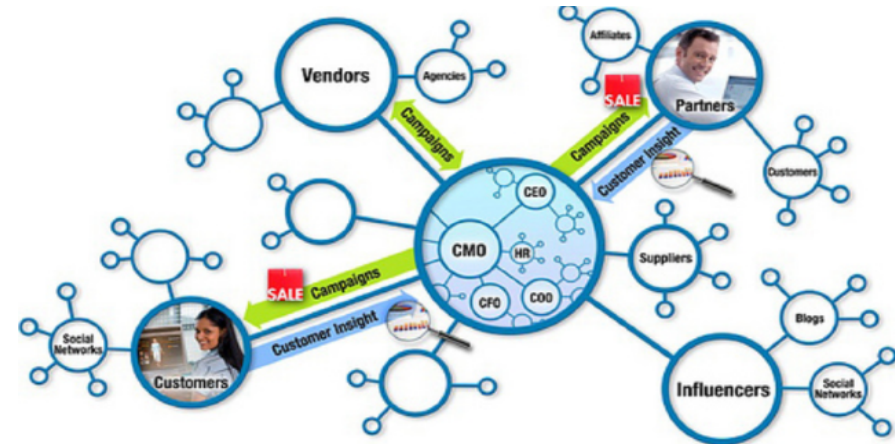
## Recommendation



### Recommended for you

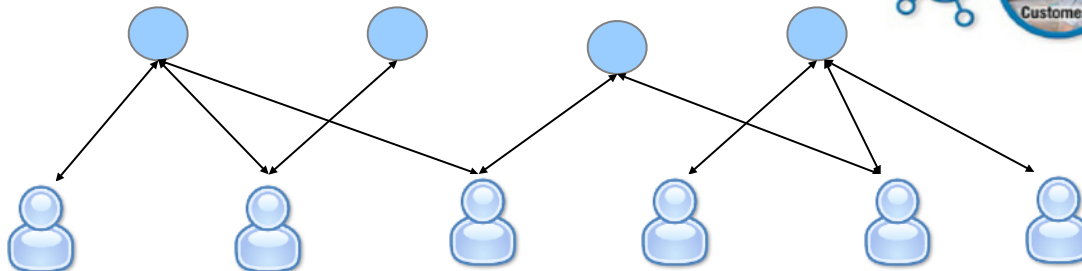


See more Recommendations



item

user



## Graph Visualizations

Communities

Graph Search

Network Info Flow

Bayesian Networks

Centralities

Graph Query

Shortest Paths

Latent Net

Ego Net Features

Graph Matching

Graph Sampling

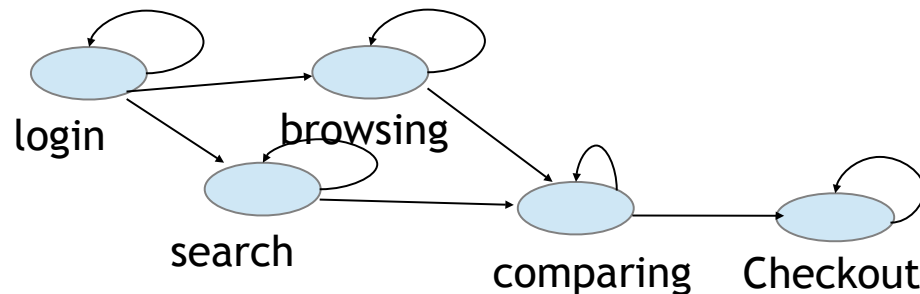
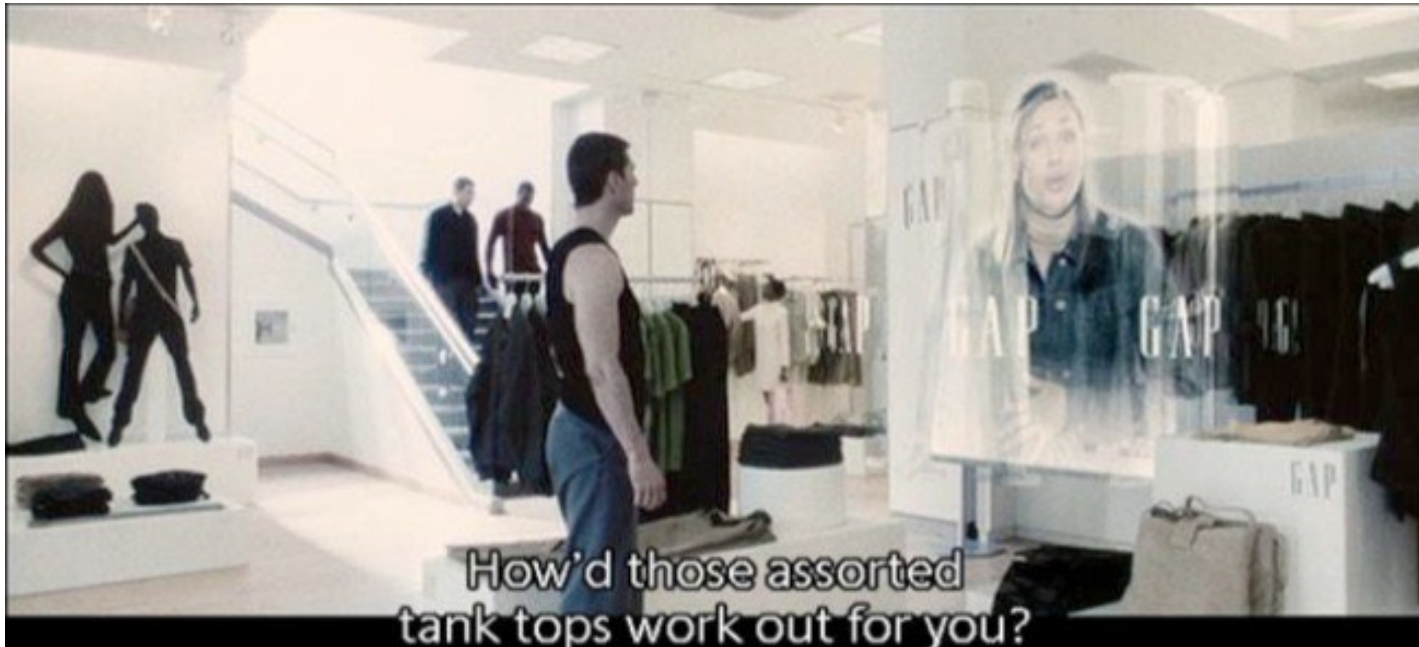
Markov Networks

## Middleware and Database

Markov  
Network

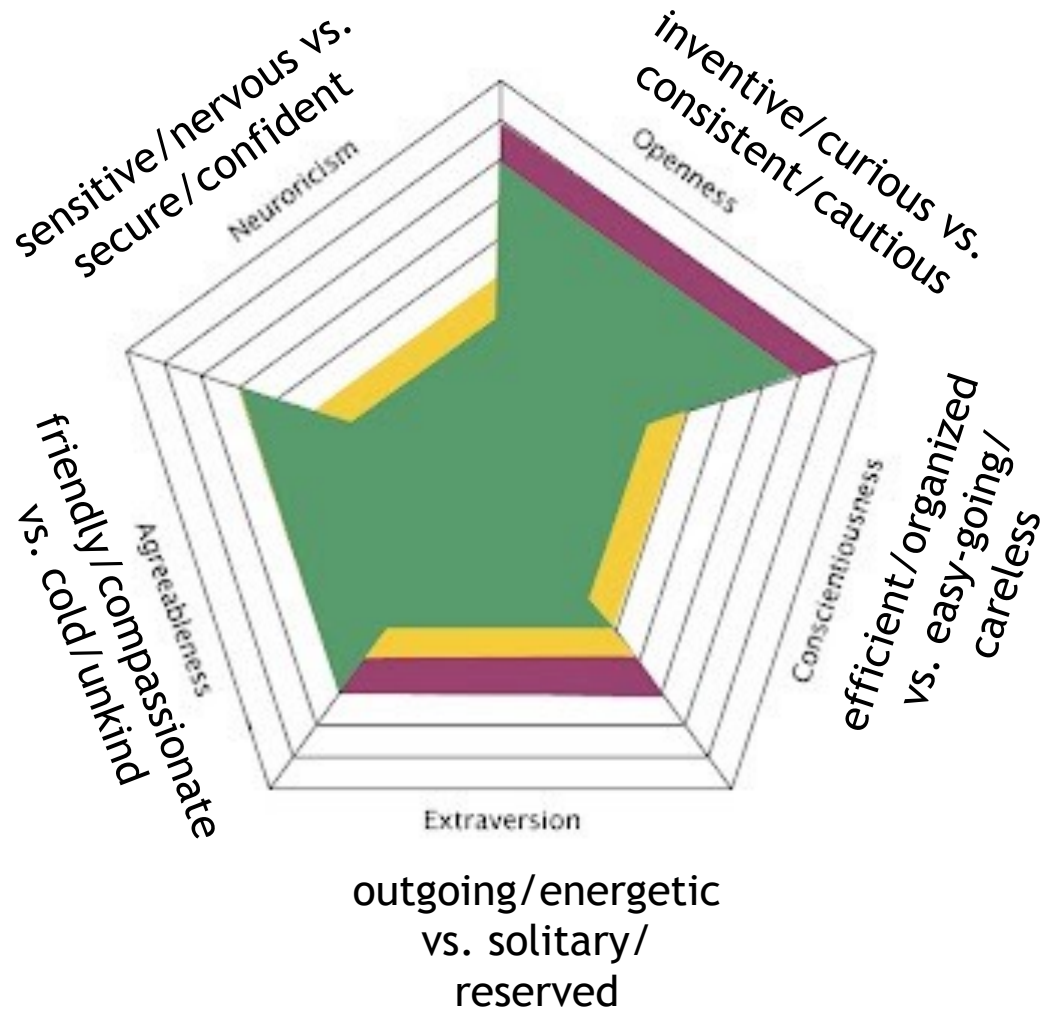
Latent  
Network

Bayesian  
Network

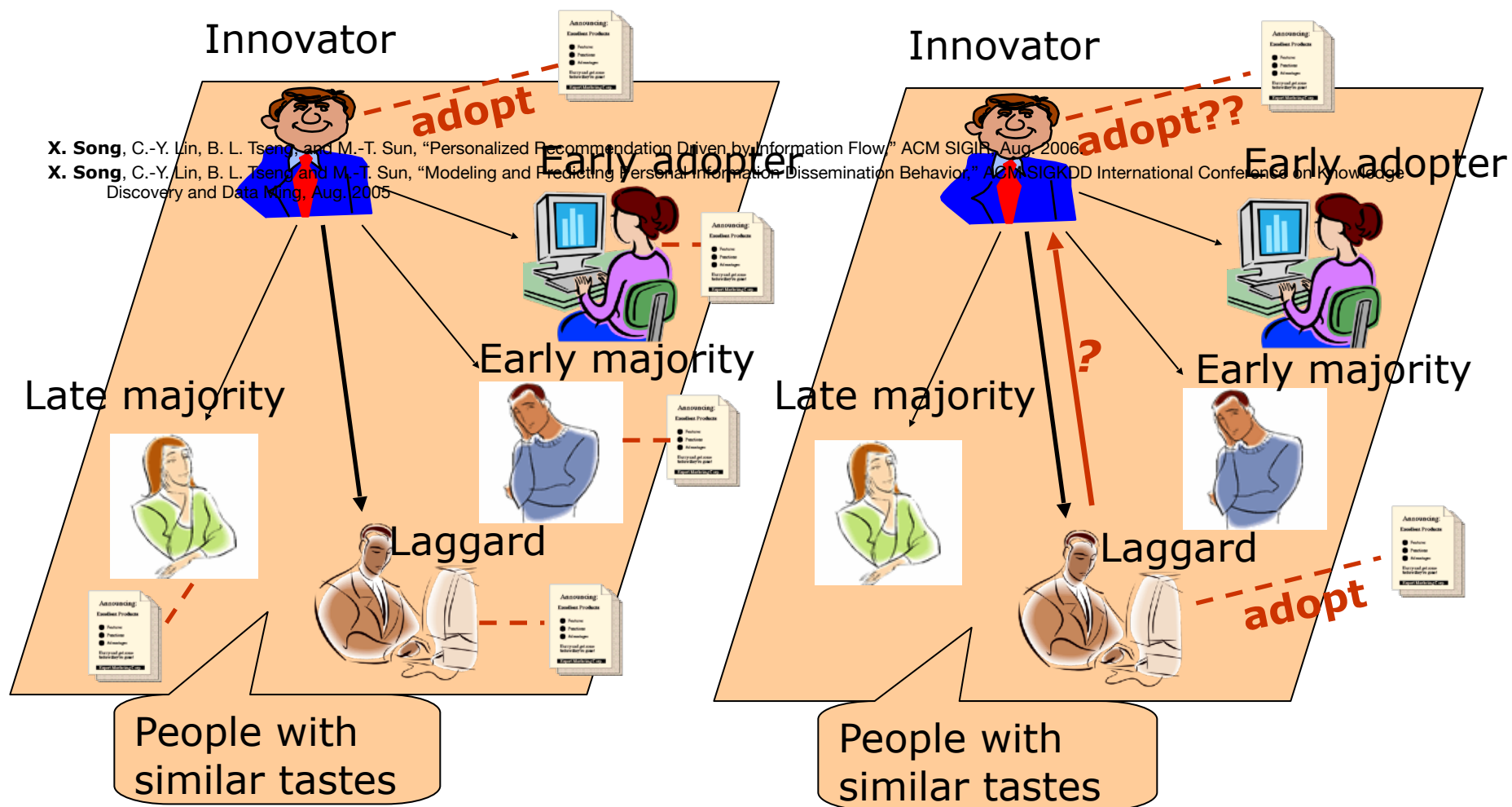


- Behavior Pattern Detection
- Help Needed Detection

## Big5 Personality (OCEAN)



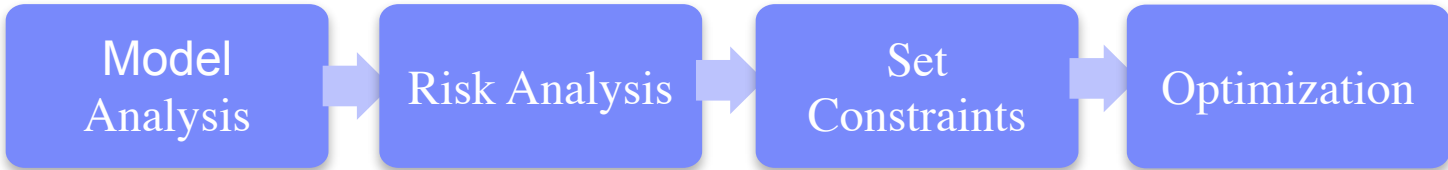




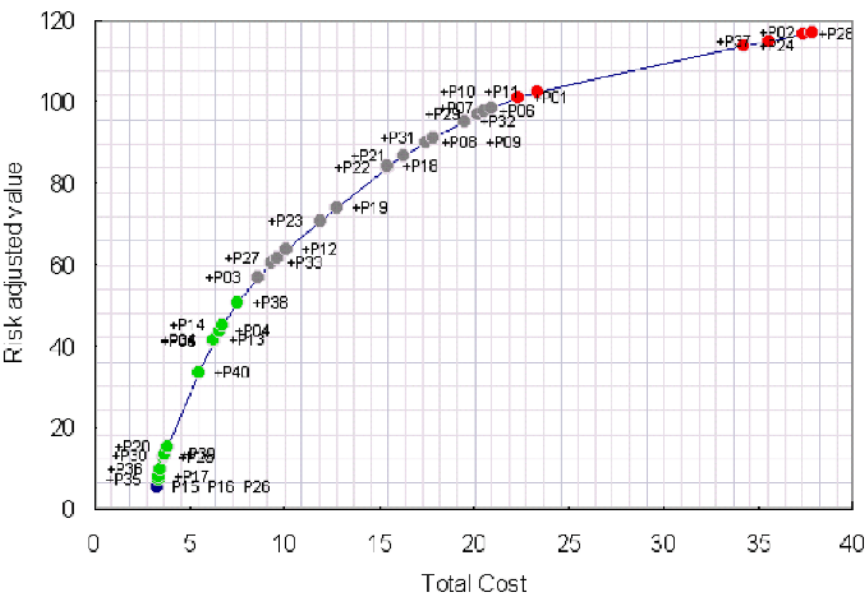
Influence is not symmetric!

- Project customer existing portfolio performance over a time period vs. suggested adjustment projected performance over the same period.
  - Show past historical similarity and simulated projection.
- Portfolio adjustment should include both conservative and aggressive bias and let the customer choose change or no change to his portfolio.
- Give customer the decision making power to make portfolio adjustment using our personalized recommendation.

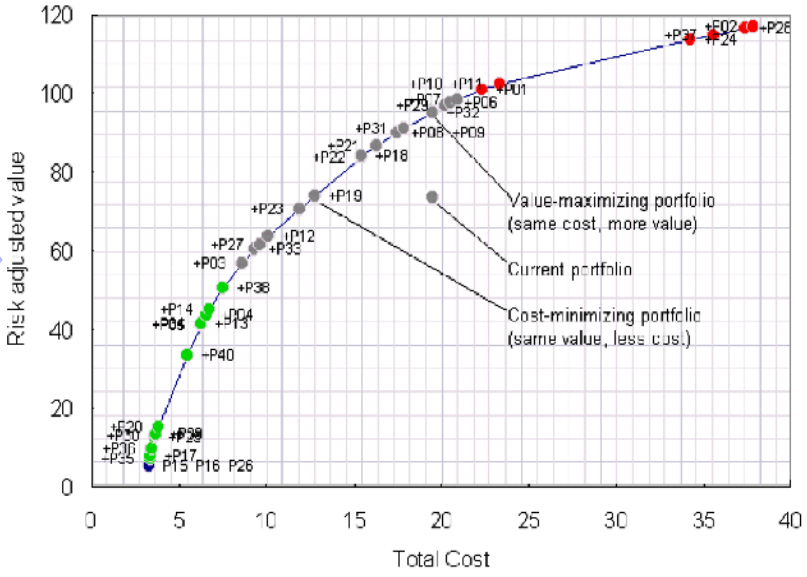
## Constrained optimized model



Risk-Profit chart



Optimized outcome



# Optimization is about Resource efficiency/utilization and allocation

| Resources        | Choices to make                               |
|------------------|---|
| Capital          | Invest, allocate                              |
| People           | Hire, assign, schedule                        |
| Equipment        | Acquire, schedule, locate                     |
| Facilities       | Locate, size, schedule, maintain              |
| Vehicles         | Acquire, route, schedule, deliver, maintain   |
| Material/Product | Acquire, allocate, produce, deliver, maintain |

- **Planning and scheduling activities**
  - Which are subject to complex operating constraints (e.g. limited resources, large volume of data, complex manufacturing or design processes)
  - With multiple business objectives to reduce time, cost, or increase KPI's such as productivity
- **While enabling**
  - Adjustment of changes in operating environment
  - What-if analysis

## Keywords:

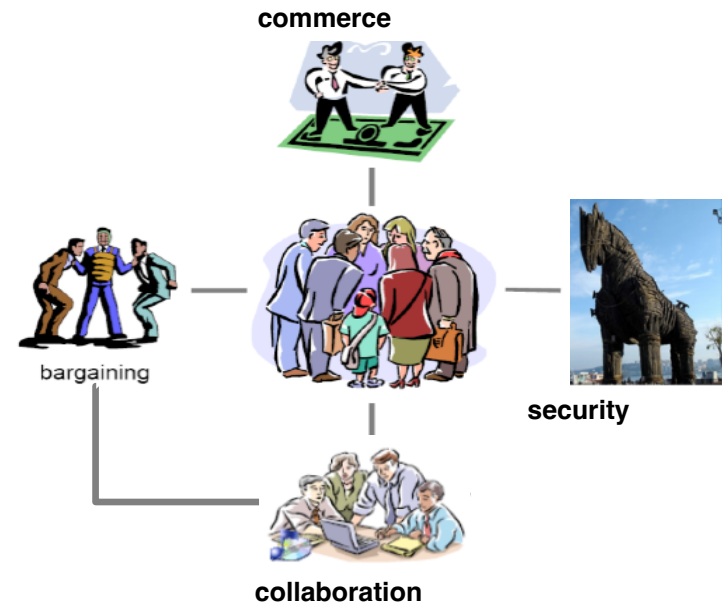
- minimize, maximize,
- how many/how much, which, when/where
- decide/choose, plan, schedule, assign, route, source, maintain, locate, trade-off

- **Issue:** Portfolio holders and managers seek maximum return from assets while limiting risks of adverse outcomes.
  - Classical formulation by Markowitz has become enriched by several factors.
  - Competitive advantage and client preferences lead fund managers to tailor portfolios to specific regional, sectoral, and other diverse preferences.
  - Novel assets have risk characteristics very different from standard stocks and bonds.
- **Scope:** Thousands of assets, hundreds of sectors, hundreds of regions. Rebalancing frequency (daily, weekly,...)
- **Decisions:** Amount of fund allocated to each asset
- **Objectives:** Minimize risk as measured by variance of portfolio return, etc
- **Requirements:**
  - Expected return at least achieves target
  - Total funds invested does not exceed amount available
  - Total funds invested per sector and/or region does not exceed limit
  - Limits on leverage

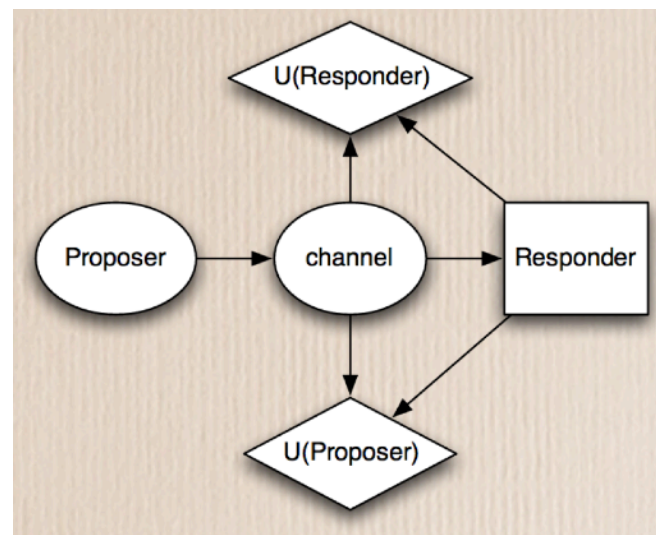
## Task 4. Bank-Customer Interaction Strategy



- Customers may not immediately accept a personalized investment strategy.
  - Customers profile may contain insufficient data (e.g. new customers) to fully capture their risk profile.
  - Customers may have their own investment strategy ideas they want to pursue.
  - Customers desired investment characteristics may be impossible to achieve
  - Customers may be willing to accept higher risk strategies than they believe.
- A personalized investment strategy may involve a number of investment stages.
  - What order should these stages be presented?
  - Can this order influence a customers risk profile?
- How can we model interaction with customers? As a multi-agent decision process, analyzed using game theory.

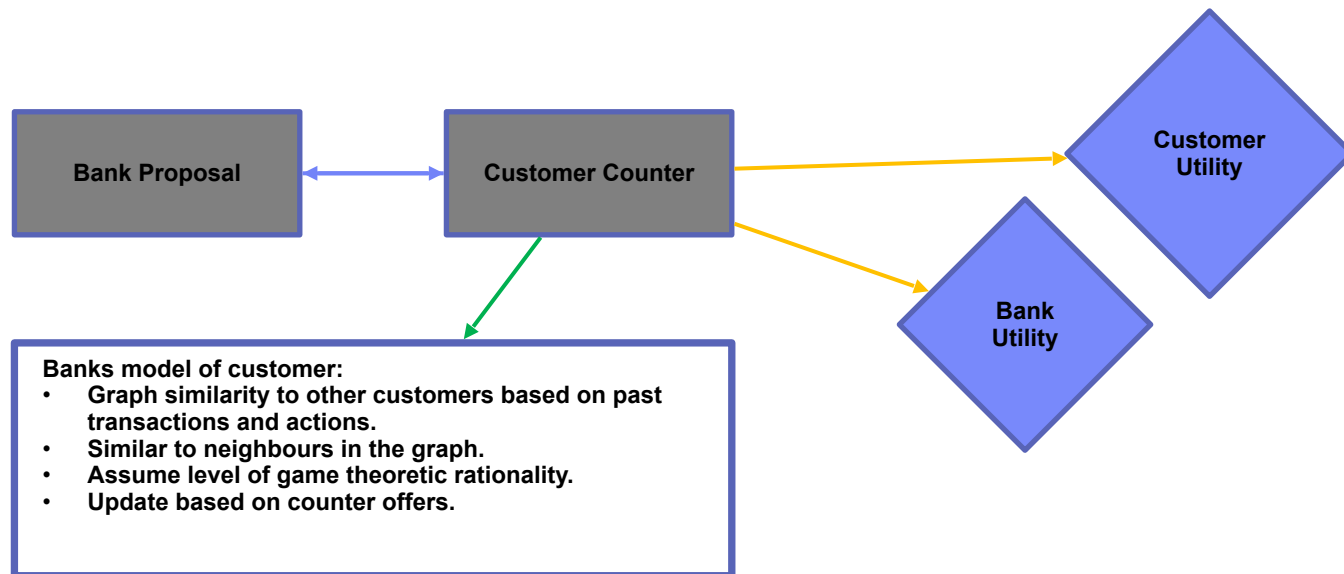


- Game theory is a system to model behavior of those in conflict, or with different goals.
  - Creates predictions about individuals using assumptions of rationality (I will make the decision that is best for me).
  - We can use an influence diagram to describe a game or decision process and solve it (e.g. using game theory).
- 
- Here we have an influence diagram representing the ultimatum game between the Proposer (i.e. a bank) and a Responder (i.e. a customer).



<http://www.eecs.harvard.edu/~gal/tutorial4perPage.pdf>

- Using influence diagrams, we can model the process of suggesting investment strategies to customers. This involves decision processes with potentially multiple interactions between a bank and each client.
- We assume that the bank and the customer each have their own financial investment characteristics that they find desirable.
- Game theoretic decision processes to settle on an investment strategy that both find acceptable, involving offers and counter offers.





AI Trader





Anita

Graphen Artificial Intelligence Traders

[Home](#) [Demo](#) [Technologies](#) [Login](#)

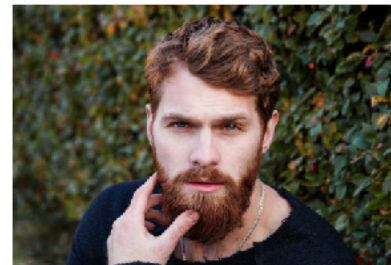
Anita avatars are earning: **\$1,501.65**



ANITA-324658  
PER \$1,000 EARN: **\$82.24**



ANITA-253758  
PER \$1,000 EARN: **\$27.04**



ANITA-247917  
PER \$1,000 EARN: **\$291.07**



ANITA-428339  
PER \$1,000 EARN: **\$55.16**



ANITA-164762  
PER \$1,000 EARN: **\$33.69**



ANITA-450214  
PER \$1,000 EARN: **\$161.56**



ANITA-247502  
PER \$1,000 EARN: **\$51.40**



ANITA-267139  
PER \$1,000 EARN: **\$456.80**



Anita  
Graphen Artificial Intelligence Traders

Home ForeignExchange Stocks Bonds



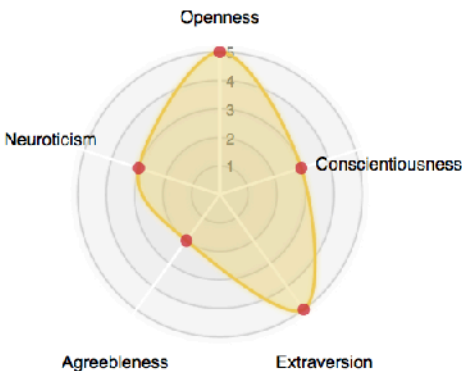
## Anita 267139

-- an Adventurous AI Trader

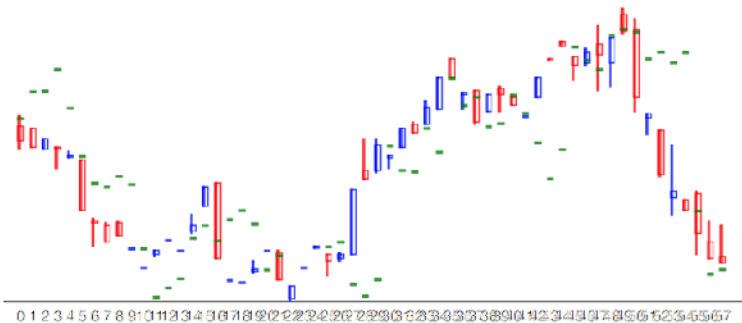
*Specialized at: EUR-USD*

*Knowledgeable of: Oil, Gold and Twitter*

*Strategy Learning Frequency at: 2.0 hours*



Original: \$1,000.00, Current: \$1,404.50, Performance: *Gain \$404.50*



### Activities

| Time                | Action       | Cash         | Unit    | Balance    |
|---------------------|--------------|--------------|---------|------------|
| 2017-10-12 13:45:05 | Sell 50,000  | \$1,404.50   | 0       | \$1,404.50 |
| 2017-10-12 12:57:25 | Buy 100,000  | \$-57,792.00 | 50,000  | \$1,386.50 |
| 2017-10-12 11:19:10 | Sell 100,000 | \$60,577.00  | -50,000 | \$1,372.00 |
| 2017-10-12 11:11:55 | Buy 100,000  | \$-57,822.00 | 50,000  | \$1,366.00 |
| 2017-10-12 09:08:05 | Sell 100,000 | \$60,566.00  | -50,000 | \$1,310.00 |
| 2017-10-12 08:34:40 | Buy 100,000  | \$-57,935.00 | 50,000  | \$1,287.50 |

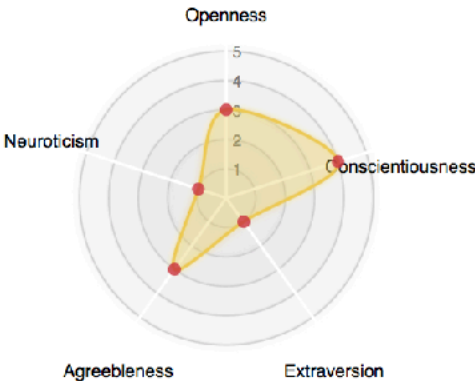




## Anita 247502

-- an Independent AI Trader

*Specialized at: EUR-USD*  
*Knowledgeable of: FX, Gold and Twitter*  
*Strategy Learning Frequency at: 100.0 days*



Original: \$1,000.00, Current: \$1,119.50, Performance: Gain \$119.50



### Activities

| Time                | Action       | Cash         | Unit    | Balance    |
|---------------------|--------------|--------------|---------|------------|
| 2017-10-12 14:58:00 | Buy 50,000   | \$1,119.50   | 0       | \$1,119.50 |
| 2017-10-12 13:56:35 | Sell 100,000 | \$60,304.00  | -50,000 | \$1,048.50 |
| 2017-10-12 11:51:25 | Buy 100,000  | \$-58,196.00 | 50,000  | \$1,012.00 |
| 2017-10-12 10:56:10 | Sell 100,000 | \$60,232.00  | -50,000 | \$992.50   |
| 2017-10-11 16:46:45 | Buy 100,000  | \$-58,236.00 | 50,000  | \$1,066.50 |
| 2017-10-11 15:13:20 | Sell 100,000 | \$60,382.00  | -50,000 | \$1,065.00 |

- Goal

- Typical portfolio allocation focuses on risk tolerance of investors which measures their ability to take risk. While investors CAN take risk, they may not be WILLING to do so. One's persona better reflects their willingness to bear risk.

- Features

- OCEAN Personality Inference and Stock recommendation from portfolio and trading history
- Portfolio and trading history rated in 6 quantifiable aspects by Market Intelligence
- Personality driven trading style classification and suggestion
- OCEAN Personality
  - Openness
  - Conscientiousness
  - Extraversion
  - Agreeableness
  - Neuroticism
- Trader Type
  - Careful Investor - John Bogle
  - Patient Investor - Warren Buffett
  - Value Investor - Benjamin Graham
  - Etc.

**Date:** December 15, 2023 **Time:** 4 - 6 pm (EE Conference Room), 7 - 9:30pm (Mudd 833)

**Format:** Remote presentation through Zoom, including on-campus and CVN students

**Each group:** ~15 mins, including demo and group QA; Preparing no more than 15 slides; stay in the whole session you're presenting

**Please be clear on these key grading points:**

**Goal & Novelty** — Existing 'challenge' or new ideas?

**Data** — 3Vs? New dataset or Existing dataset? Difficulties to process/gather data

**Technology** — methodologies and algorithms;

**System** — an overview of implemented system. What was implemented?

**Demo** — the quality of the system demo

**Materials to submit before presentation:** Project information and Slides for the course external website. (URL will be posted later).

**Deadline: December 22, 2023**

**Final Project Report (IEEE double column style, less than 10 pages)**

**Source Codes**

**Final Project Introduction Video on YouTube (no more than 10 mins).**

URL will be posted later.

- **Title, Author(s)**
- **Abstract:** Briefly describe your problem, approach, and key results.
- **Introduction (5%):** Describe and define the problem you are working on. Why is it important? Include an overview of your methods and results.
- **Related Work (5%):** Discuss published works or approaches that are related to your project. What's the benefit or drawback of the previous works? What kind of problems have they solved? How is your approach similar or different from others?
- **Data (10%):** Describe the data you are working with for your project. What type of data is it? Where did it come from? How much data are you working with? Did you have to do any preprocessing, filtering, feature engineering or other special treatment to use this data in your project?
- **Methods (25%):** Discuss your approach for solving the problems that you set up in the introduction. Why is your approach the right thing to do? Did you consider alternative approaches? Have you tried some methods that didn't work out? It may be helpful to include figures, diagrams, or tables to describe your method or compare it with other methods.
- **System Overview (25%):** Describe the software architecture and tech stacks of your application. Discuss potential bottlenecks and improvements that could be made. Mention the software packages that you used. Mention how to use your application. You could provide screenshots to your application
- **Experiments (20%):** Discuss the experiments that you performed to demonstrate your approach solves the problem. The experiments will vary depending on the project, but you might compare with previously methods, determine the impact of the components of your system, experiment with different hyper-parameters, architectures, or algorithms, use visualization techniques to gain insight of how your model works, etc. Graphs, tables, and figures are highly recommended to be included to illustrate your experimental results.
- **Conclusion (5%):** Summarize your key results. What have you learned? What problems have you discovered and solved? Suggest ideas for future extensions or new applications.
- **Writing / Formatting (5%)** Is your paper clearly written and nicely formatted?