



Graph & Machine Learning for Enhanced Analytics

Brandy A. Freitas, Senior Data Scientist



Agenda



**Introduction to
Graph Databases**



**Real World Examples:
Jim Kilgore, Principal, ICC**



**Machine Learning,
Predictive Analytics**

What is a Graph Database?





Graph databases: 2 examples

Sales planning use case: bring existing connections to light



“Succeeding in business is all about making connections”

- Sir Richard Branson

Goal

- Identifying relationships across the client’s organization and their prospect organizations
- Create a knowledgebase asset that our client can continue to build upon

What is the data set constructed for the project?

Client organization

- Top level executives, Board of Directors, sales team
- Customers

External

- Affinity organizations
- Partner organizations

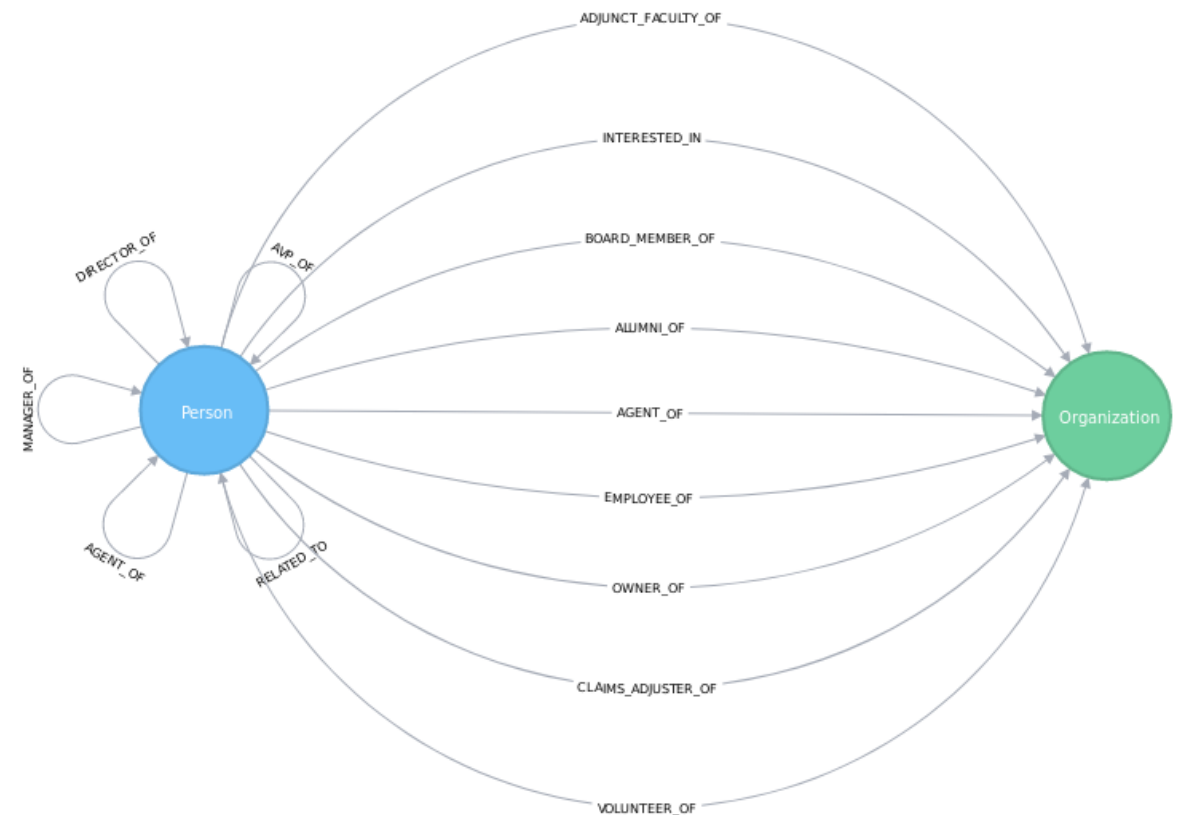
Prospect

- Key executives and board members

LinkedIn profiles (for all of the above)

- Alumni relationships, Board relationships, Other interests

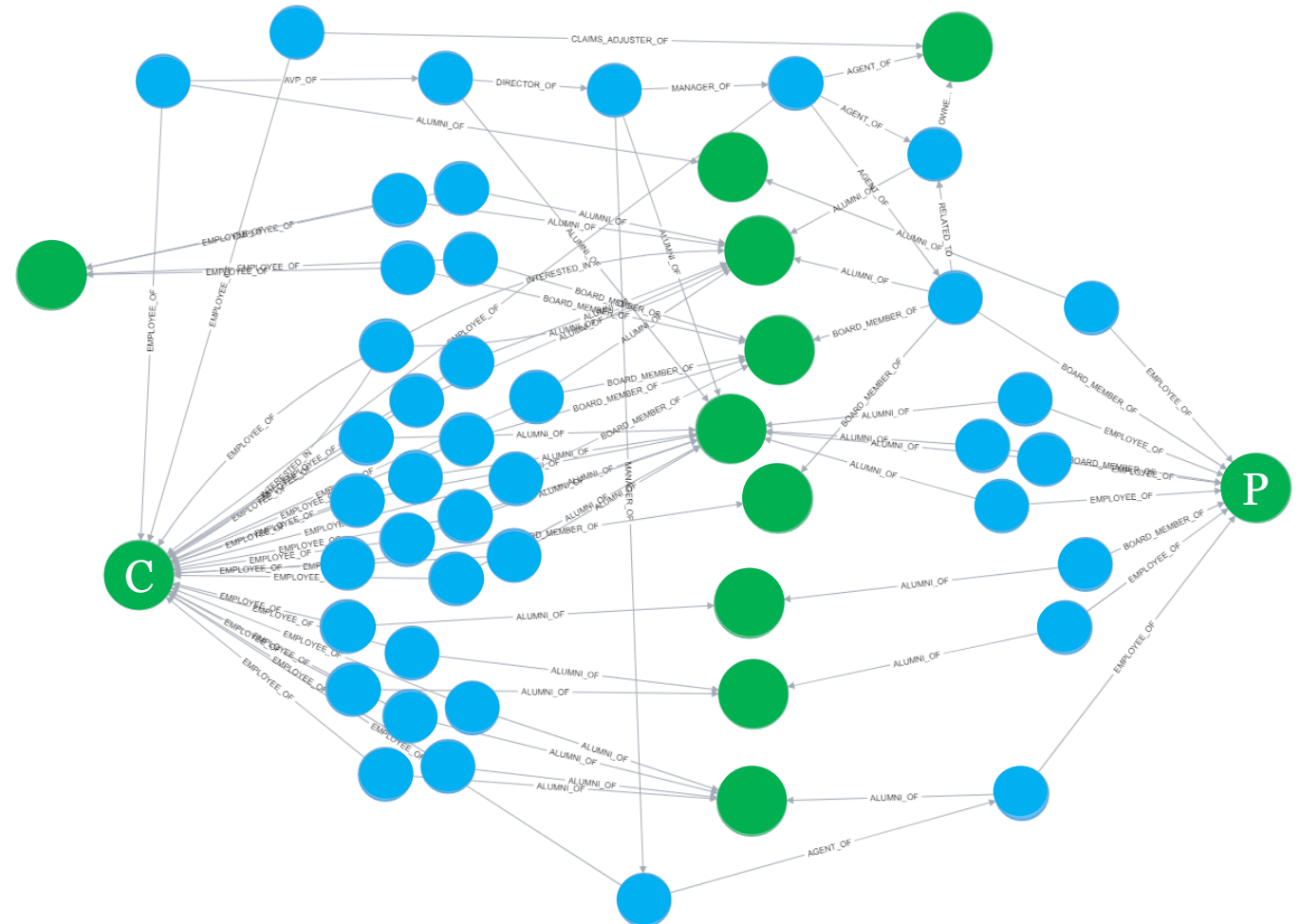
The data model



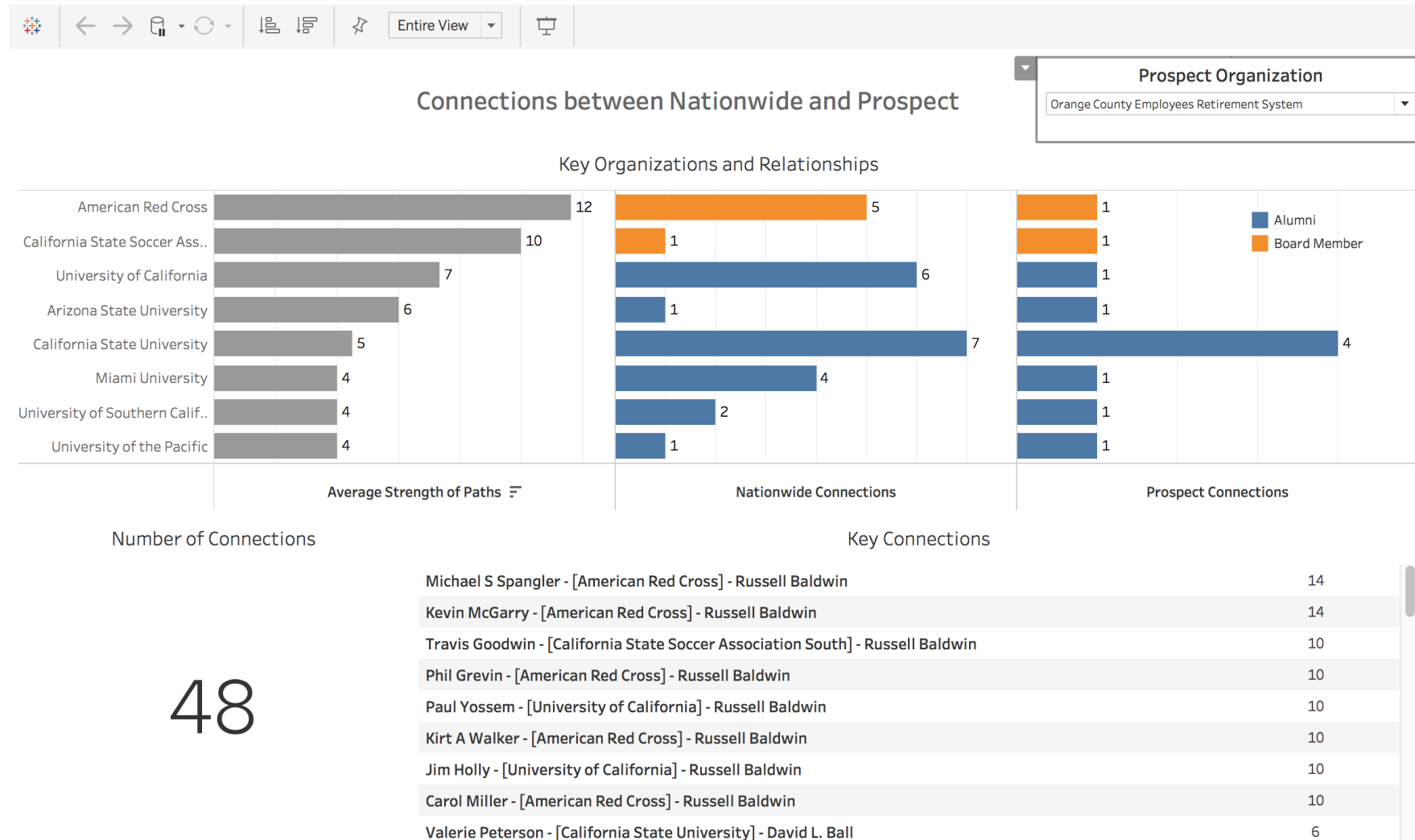
What connections did we identify?

Relationships between the client and their prospect

- 48 relationships identified



How a Salesperson might access via a Tableau dashboard



48

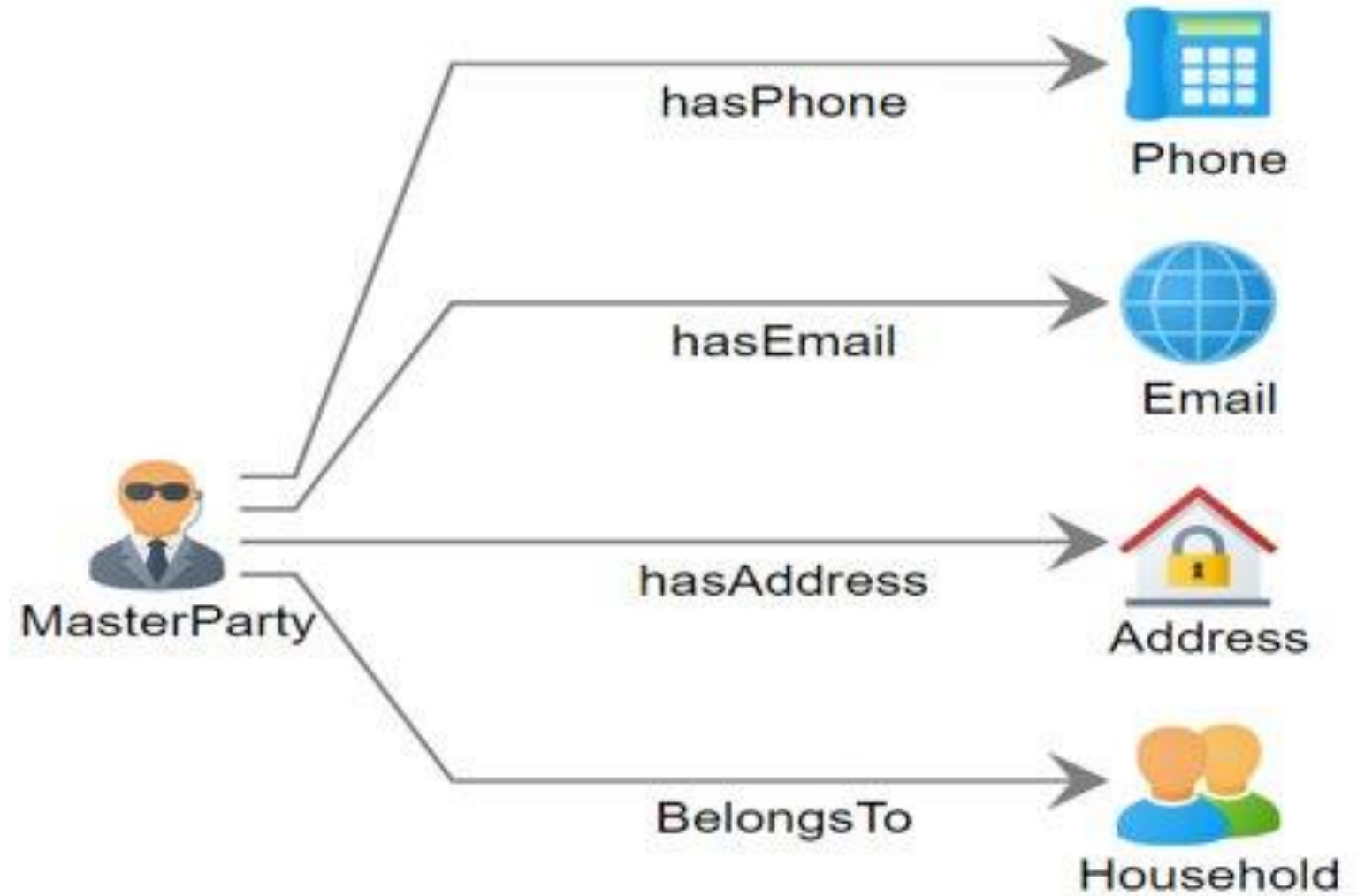
Customer analytics use case: making probabilistic matches across customer records

Order #	Channel	Name	Date
123	Store 12	J. Johnson	
234	store.com	Jennifer Johnson	
345	Mail reply	Jenifer Johnson	
456	Store 22	Jen Johnson	
567	Call center	Jen Johnston	

Goal

- Consolidate customer records and transactions across web, retail stores and paper-based applications
- Create Lifetime Value calculation with more complete view of current and potential value

Customer Data Model



Graph algorithms enable setting thresholds scores

Load match rule: **Custom - Name and Party and (Address or Email,Phone)** New Save...

Group by: **PartyDeduplicationMatchKey1** Advanced ☐ Enable sliding window
☒ Sort Window size: 50
☐ Express match on: ☐ Assign collection number 0 to unique records
Initial collection number: **1**
☐ Generate data for analysis

Name and Party and (Address or Email,Phone)

- Party Name
 - Personal Name
 - PartyFirstName
 - and PartyMiddleName
 - and PartyLastName
 - and PartyMaturitySuffix
 - or PartyCompanyName
- and Address or (Email Phone)
 - Party Address
 - Street
 - PartyAddressLine1
 - and PartyAddressLine2
 - and City/State or Postal Code
 - City and State
 - PartyCity
 - and PartyState
 - or PartyPostalCode
 - and PartyCountry
 - or Email Phone

Add Parent Add Child Remove Move Up Move Down Evaluate...

Parent Options
Name: **Name and Party and (Address or Email,Phone)**
☐ Match when not true

Matching Method
☐ All true ☐ Any true ☒ Based on threshold

Missing Data
☒ Ignore blanks ☐ Count as 0 ☐ Count as 100 ☐ Compare Blanks

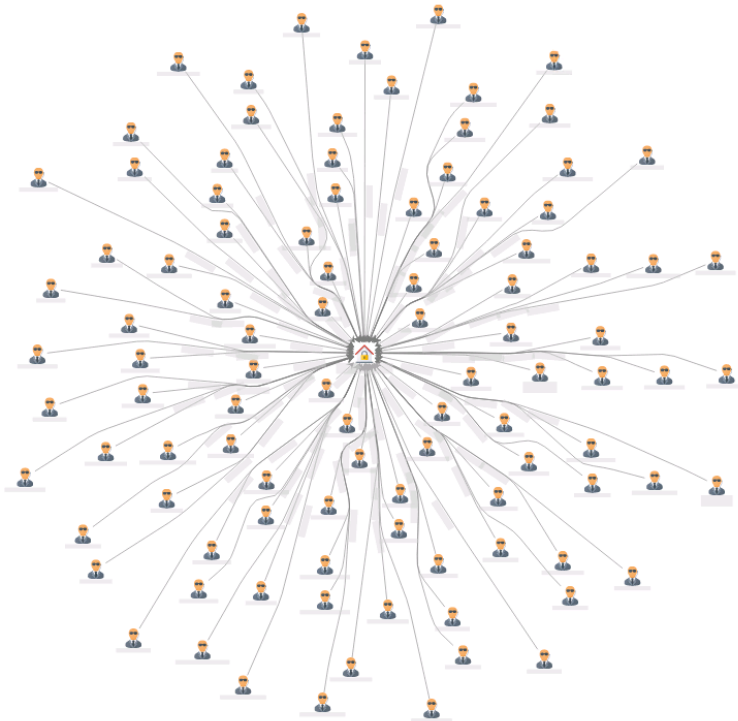
Threshold 80

Scoring method: **Average**

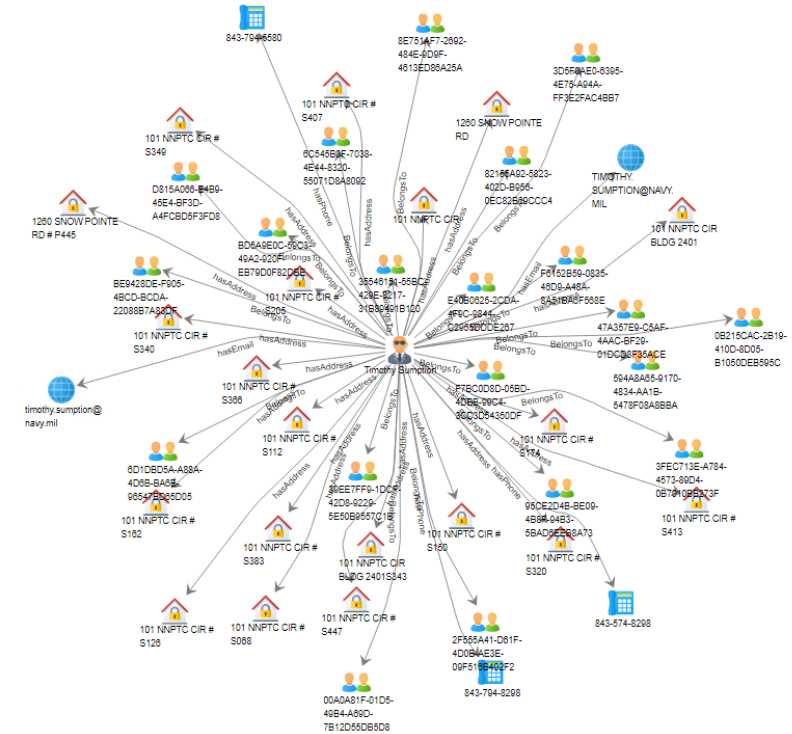
Child
Party Name
Address or (Email,Phone)

Validating rules with graph visualizations

Address with multiple Customers



Customer with multiple Addresses



Why these 2 use cases for graphs?

- The relationships drives the insight
- Retrieval of relationship is fast
- Pattern detection is needed
- The data is incomplete across key fields
- The data is inconsistent
- Flexibility to add to the graph asset

Enhanced Analytics on Graph



Machine Learning
Augmentation

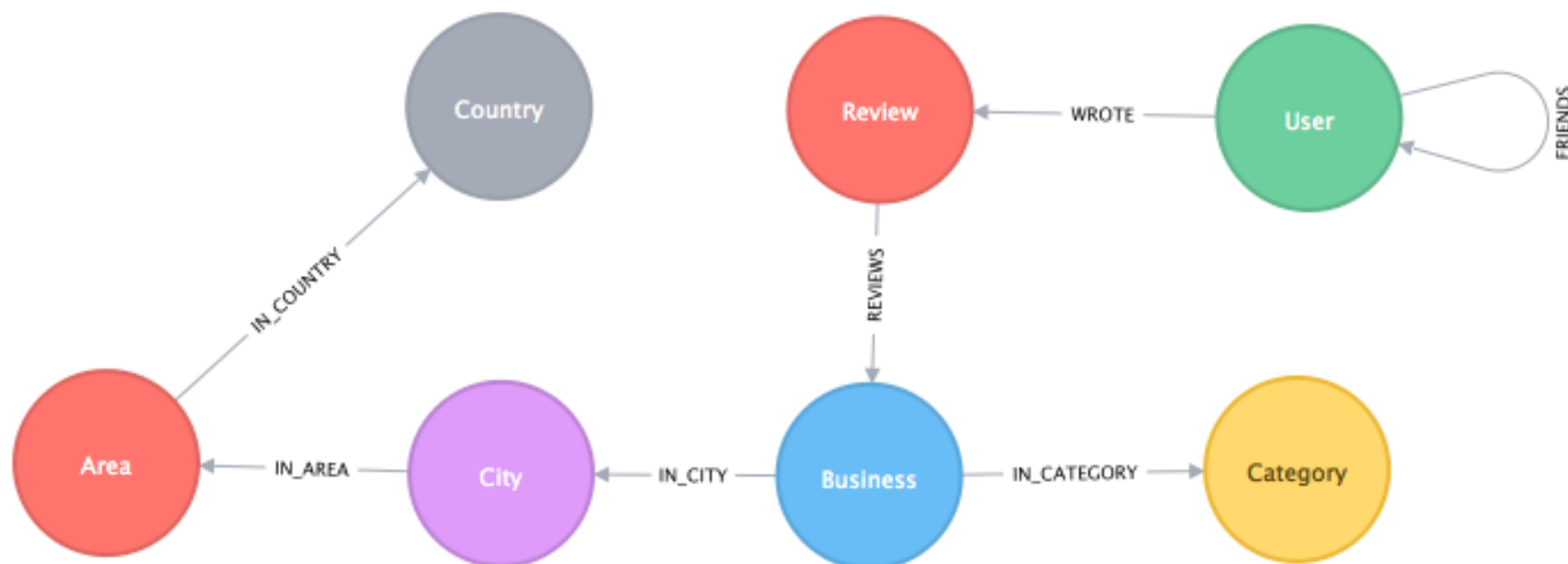


Predictive
Analytics

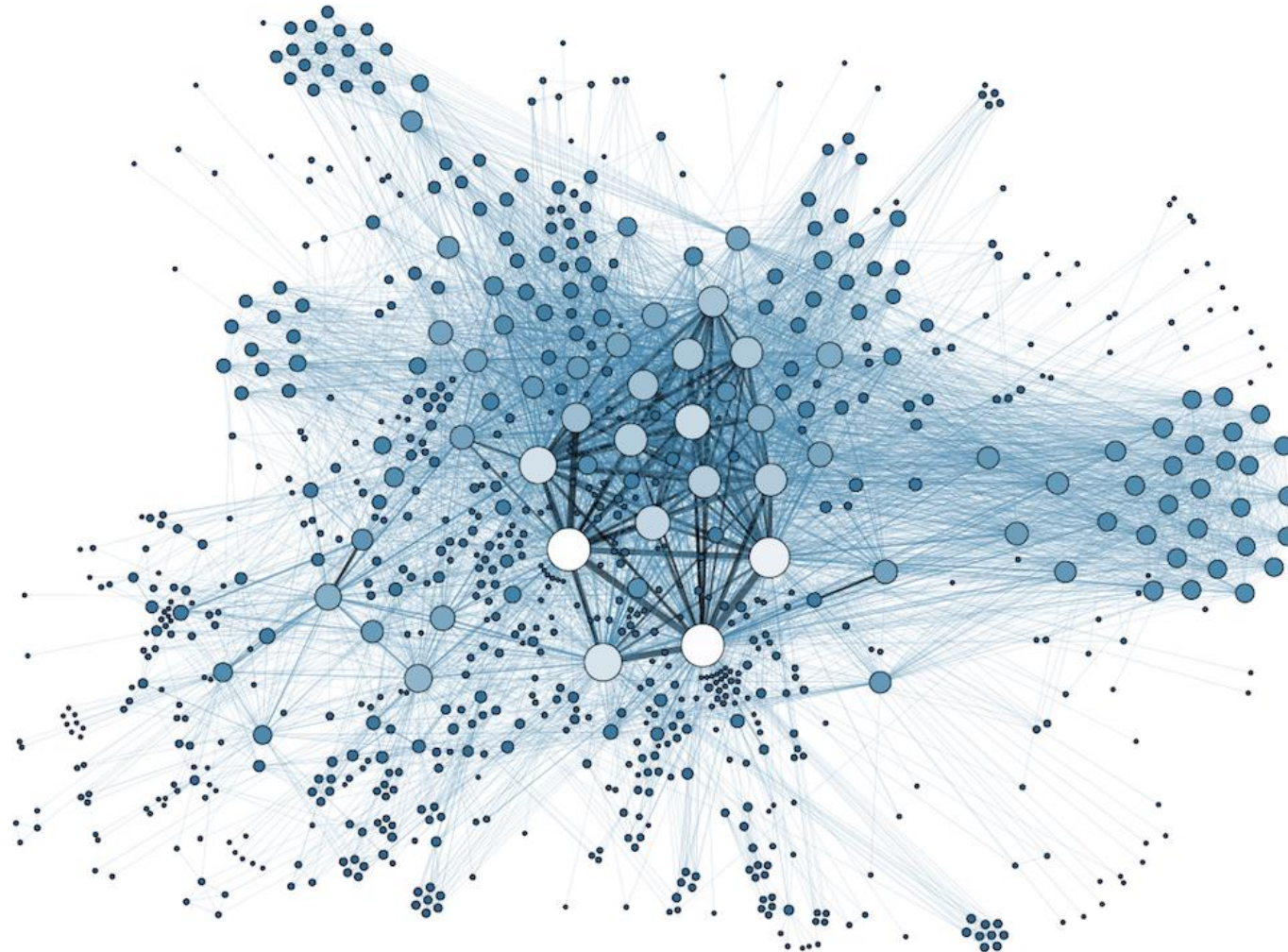


Native Graph
Algorithms

Machine Learning



Predictive Analytics



Native Graph Algorithms

