

Best Practices for Deploying a Fraud Analytics Solution

Miriam Friedel, Ph.D.
Robert Han

Headquarters

300 W. Main Street, Suite 301
Charlottesville, VA 22903
434.973.7673 | fax 434.973.7673

www.elderresearch.com
Copyright © 2017 Elder Research, Inc.

Office Locations

Arlington, VA
Linthicum, MD
Raleigh, NC

Webinar Agenda

- Introduction to Fraud Analytics
- Elements of fraud detection with case studies
 - Data Acquisition and Modality
 - Deployment
 - Organizational Maturity
- Q&A

About Elder Research

Elder Research delivers business value through customizable advanced analytics solutions that solve your most challenging problems.



20+ years
experience



150+
customers



Data science
experts



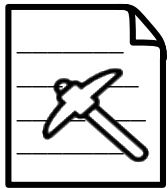
Trusted
partner

Advanced Analytics is Our Strength



Data Science and Predictive Analytics

Discovering patterns in past data that can be used to predict the outcome of future events including statistical modeling, classification & analysis, clustering, optimization & simulation, and customer segmentation



Text Mining

Understanding information stored in text documents and databases including document classification, natural language processing, information extraction and search



Data Infrastructure

Cleaning, preparing, and integrating disparate data sources and building ETL and data pipelines optimized for advanced analytics



Data Visualization

Making advanced algorithms easily accessible through 2-D & 3-D, statistical and spatial visualization

Introduction to Fraud Analytics

Defining the Problem

- Fraud analytics, as explained to a seven-year old:

People are trying to do bad things, usually to get more money. We are trying to stop them using data.

Data Triage

You and your team are trying to detect fraud.



Data Triage

If you can only examine so many people, how do you decide which ones to investigate?



Data Triage

You could search randomly...



Data Triage

You would catch some but you would miss many others



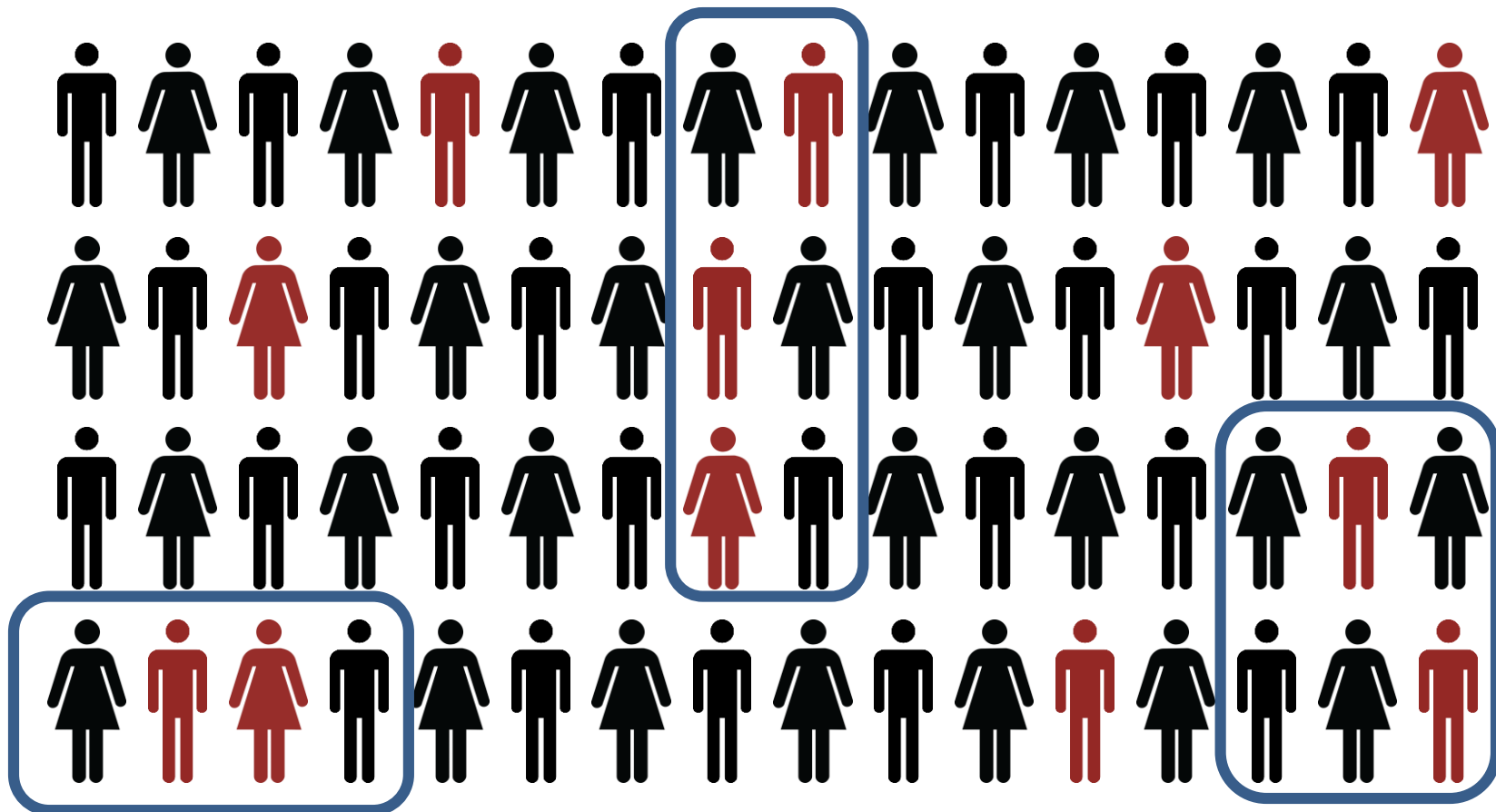
Data Triage

Or you can use a statistical model...



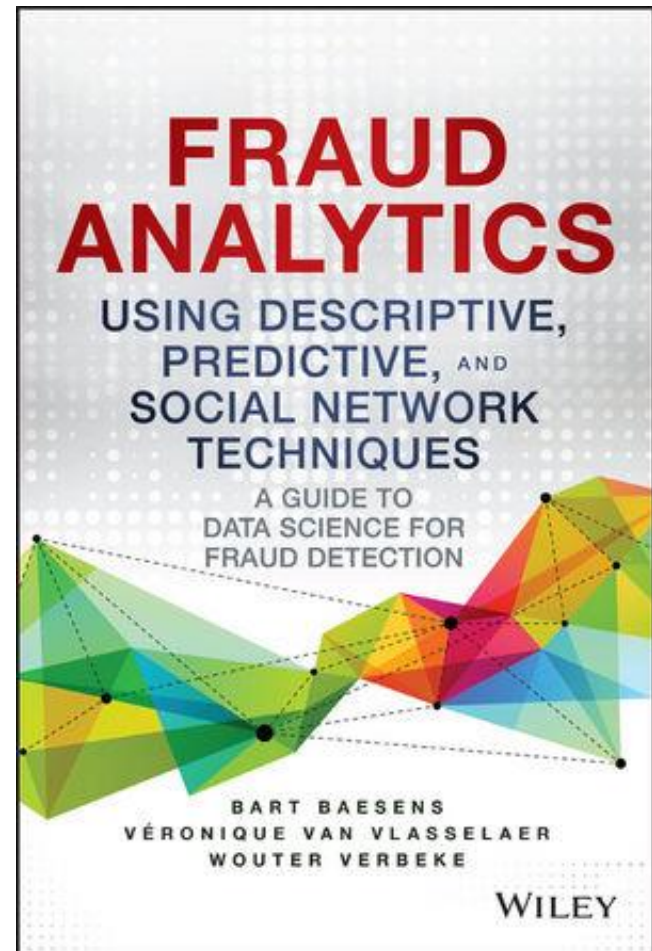
Data Triage

You would still have some misses, but you would detect more fraud, and you can measure your performance!



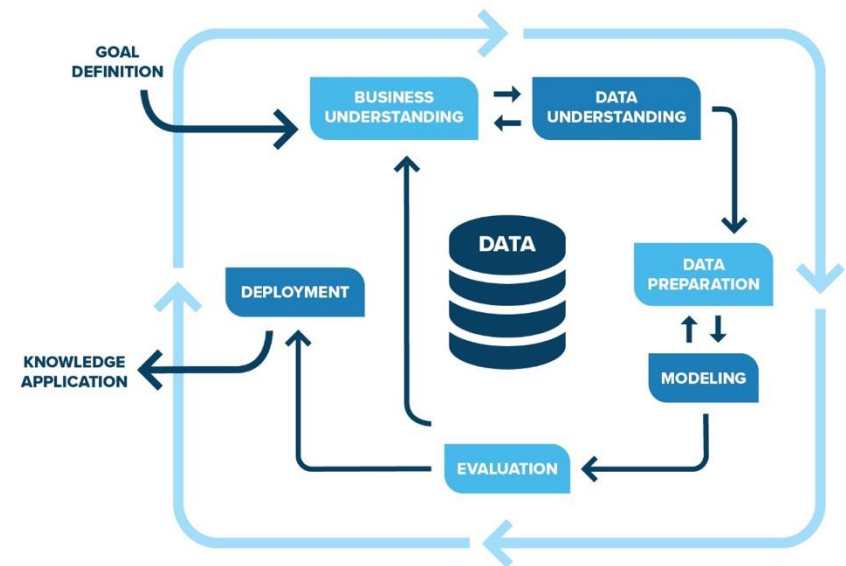
How is fraud detection different?

“Fraud is... well-
considered,
imperceptibly
concealed,
time-evolving, and
often carefully
organized crime
which appears in
many types of forms”



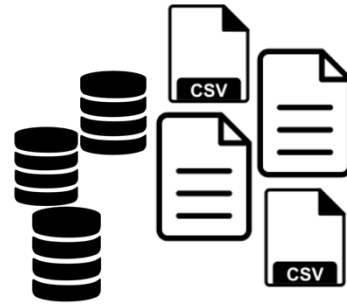
How is fraud detection the same?

- The good news is that analytics best practices also apply to fraud detection!
 - Problem context and framing
 - Data ingestion and transformation for modeling
 - Appropriate technology stack



Some Dimensions to Consider

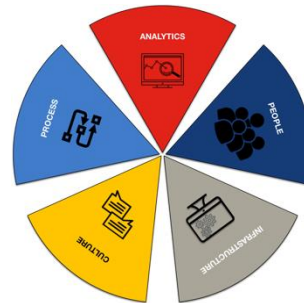
- Data Acquisition and Modality



- Deployment



- Organizational Maturity

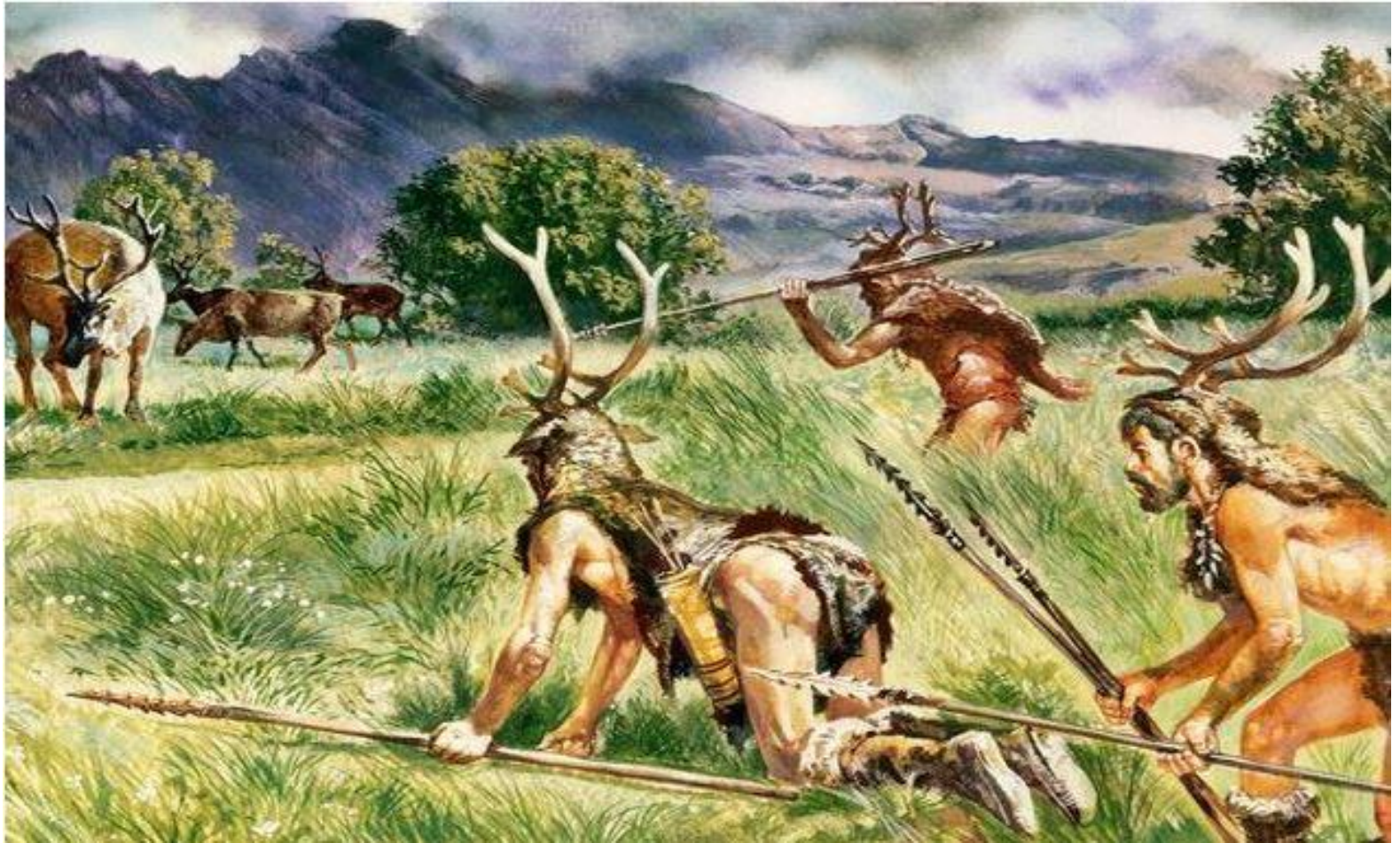


Data Acquisition and Modality

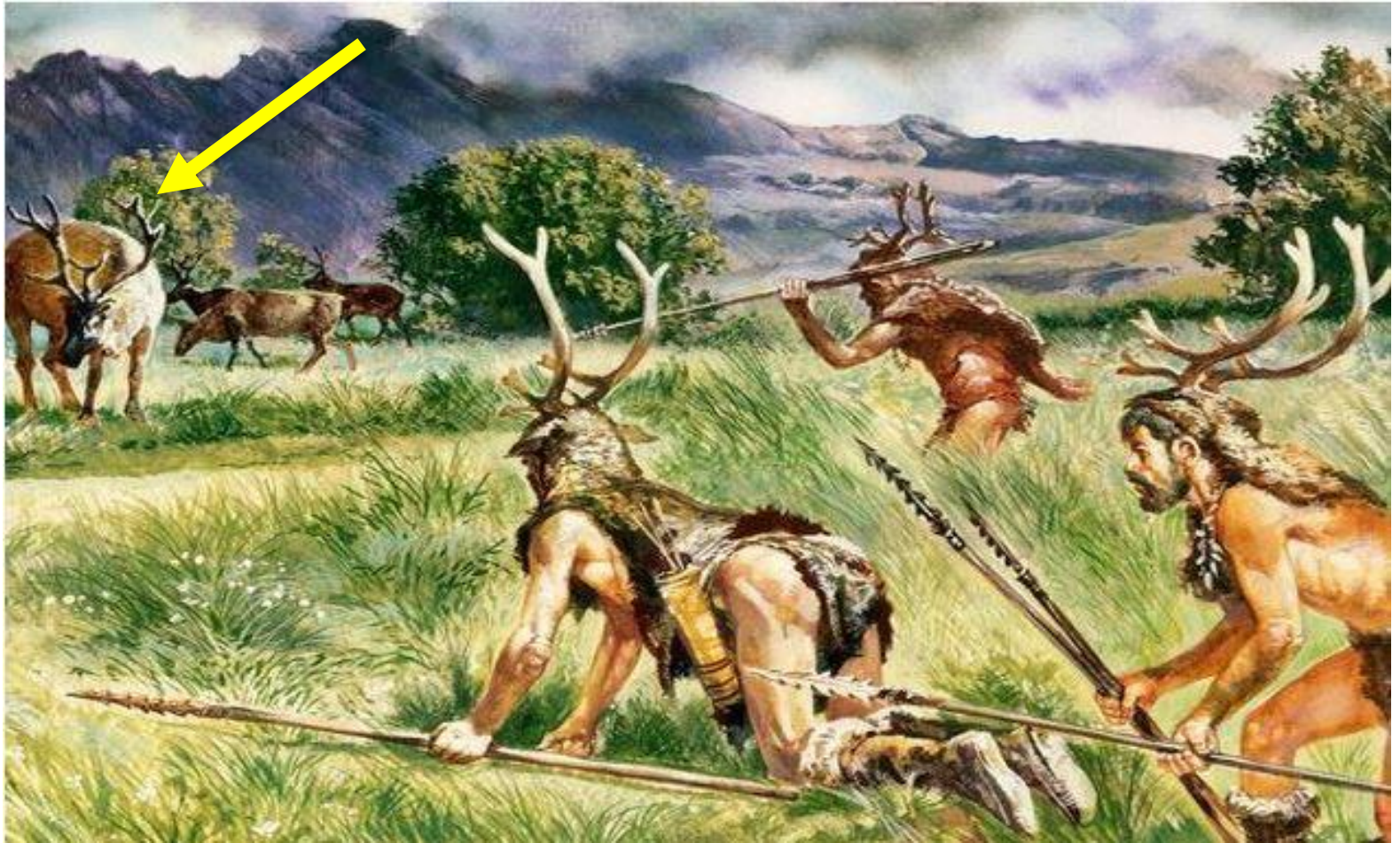
Data Hunter/Gatherer

- Data starved
 - ...Not very wide
 - ...Not very rich
 - ...Not many known cases (targets)
 - ...Not easily accessible (silos)

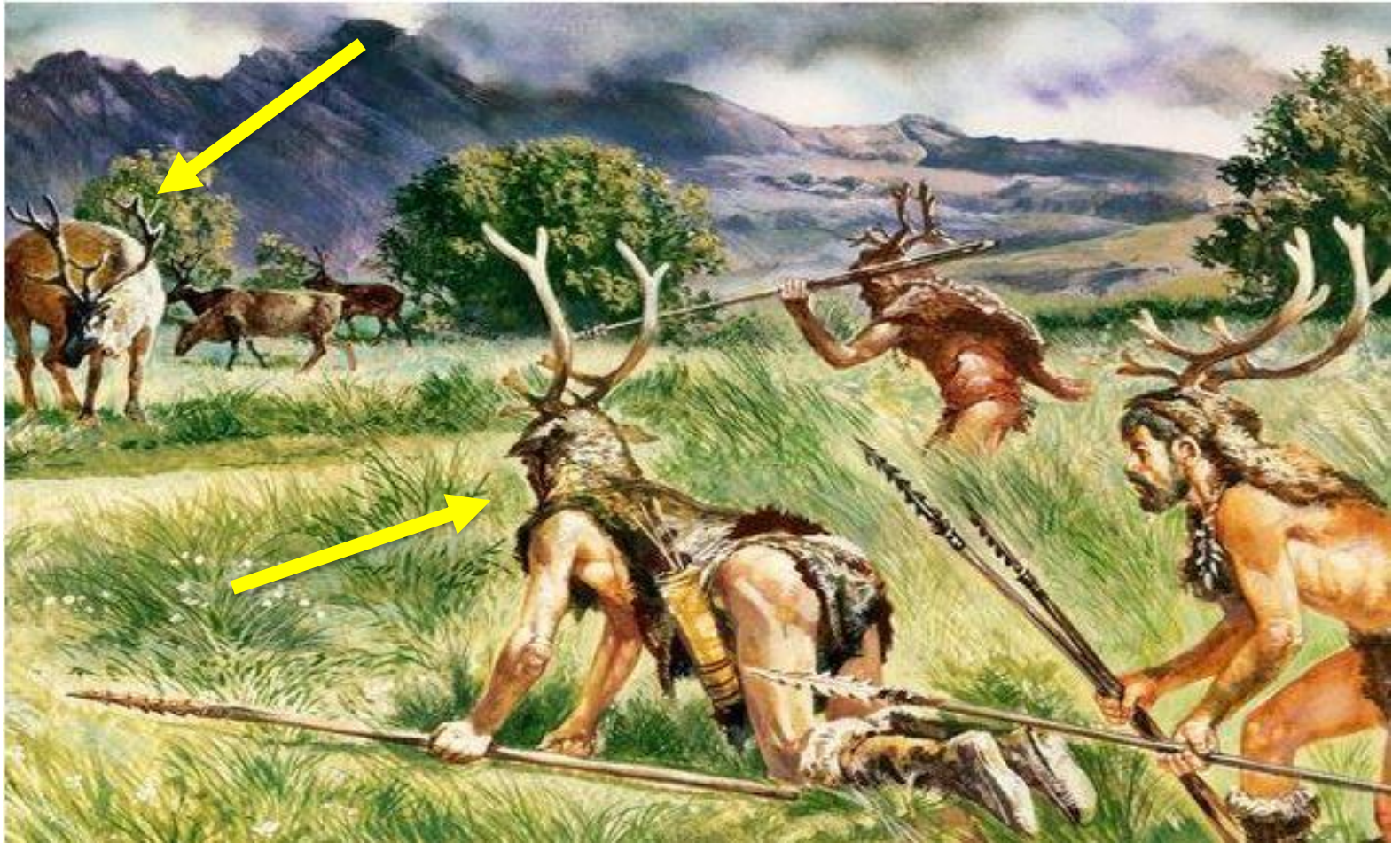
Data Hunter/Gatherer



Data Hunter/Gatherer



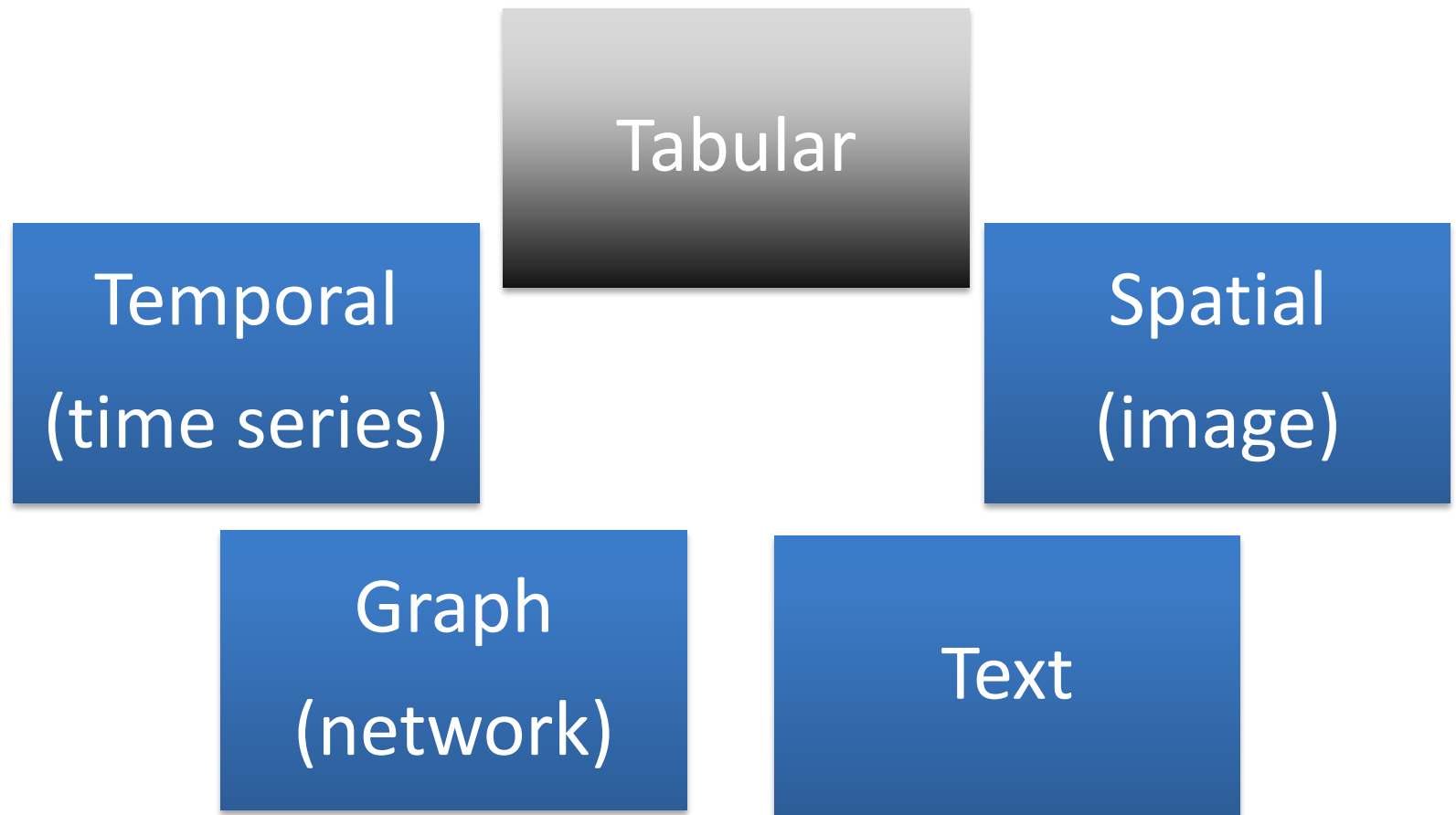
Data Hunter/Gatherer



Consider the modality of your data

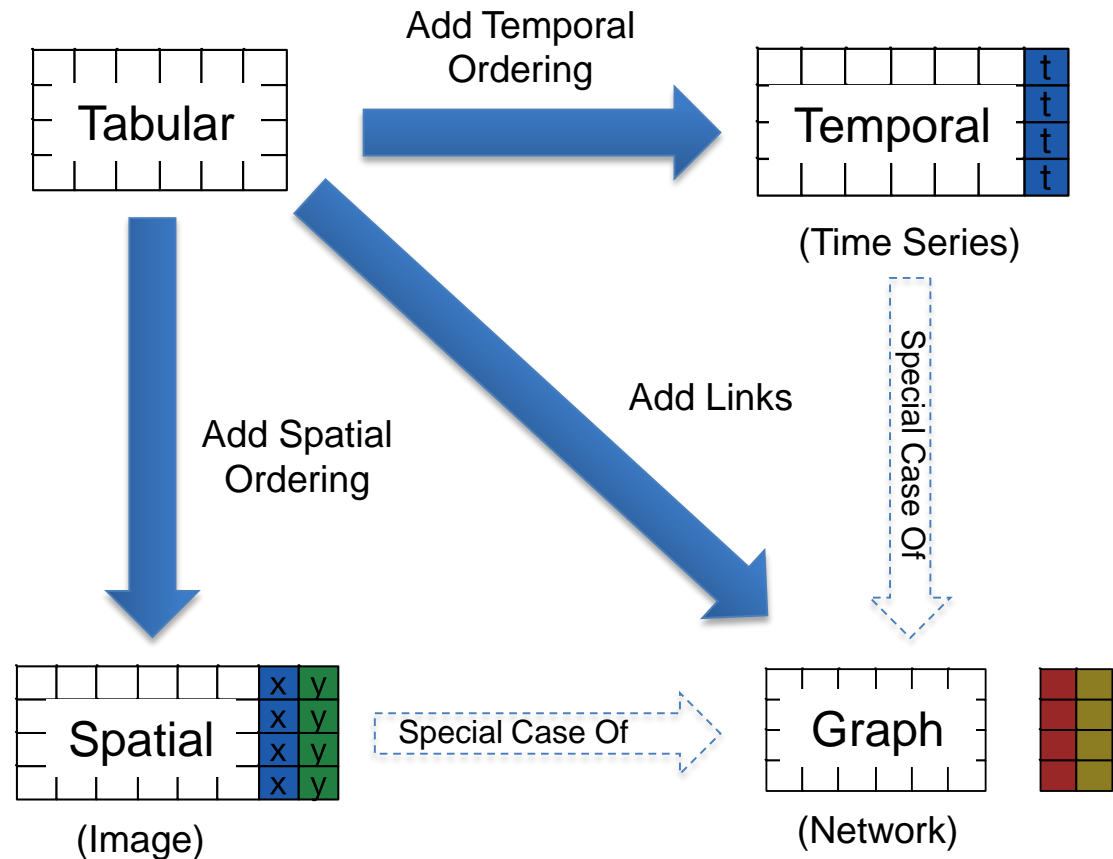
- Different ***modes*** of data lend themselves to various types of techniques

Various Data Types



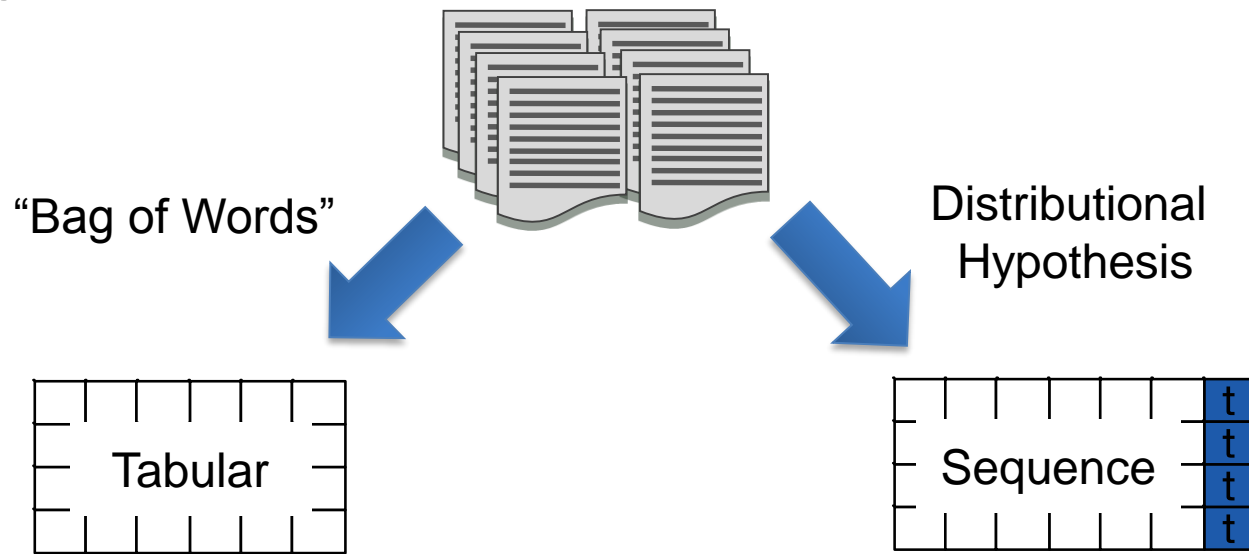
Adding Complexity

- Advanced data types each add a new type of complexity to tabular data
- They allow the representation of richer concepts...
- ... but, require special techniques to model

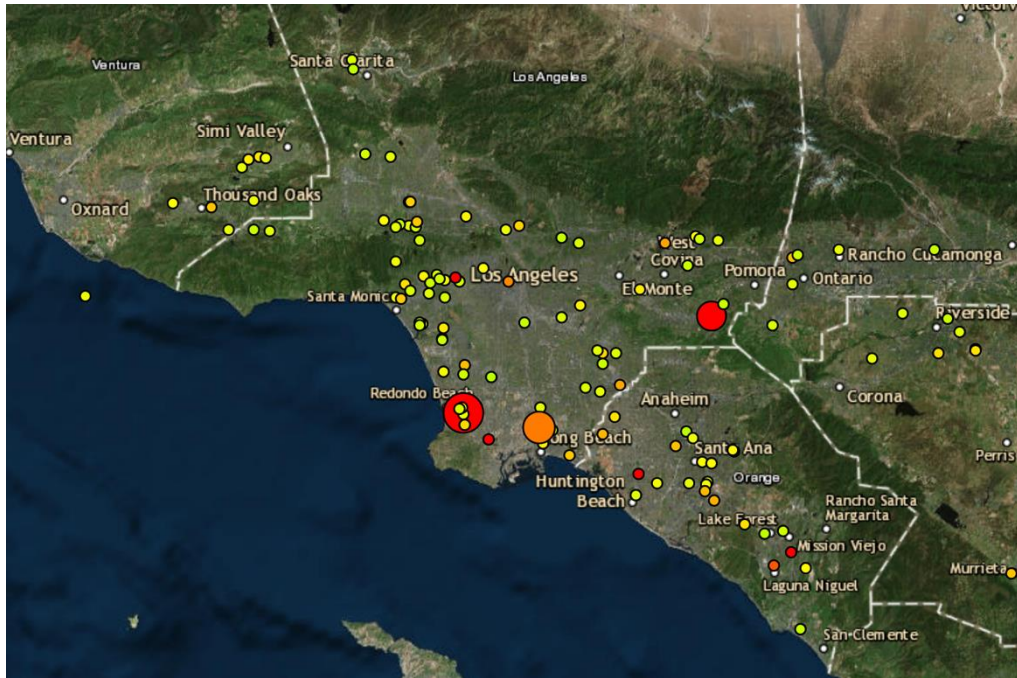


Text Data: Transformation Required

- Unable to process symbolically, must convert to numerical form
- Then apply all usual techniques for that data type

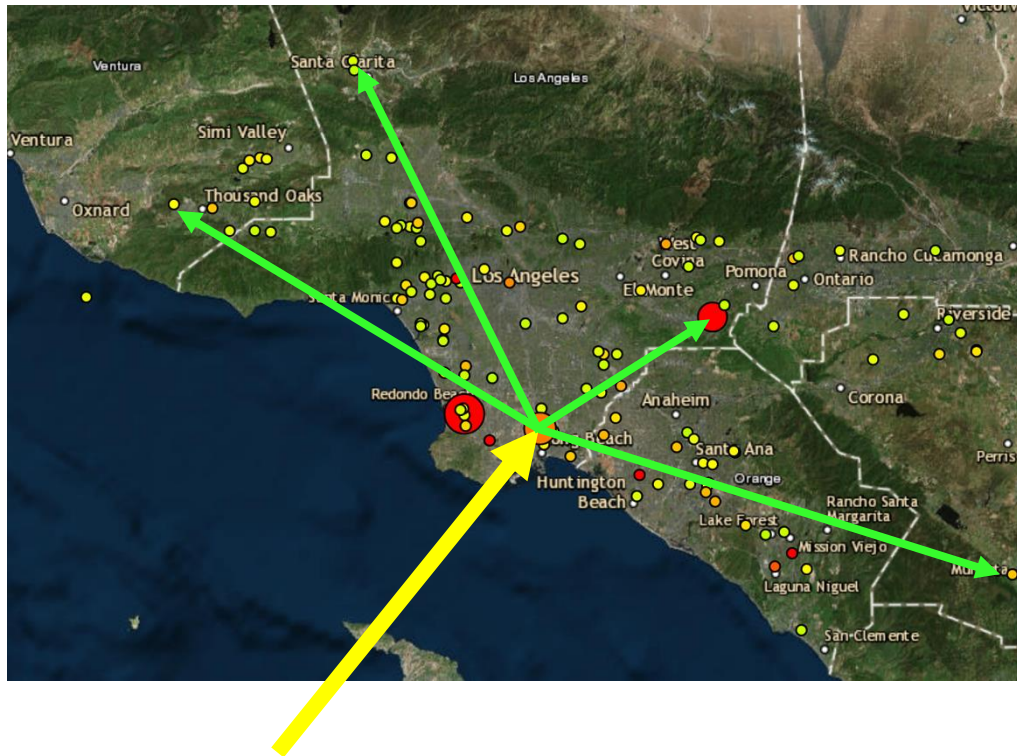


Example: Spatial Data



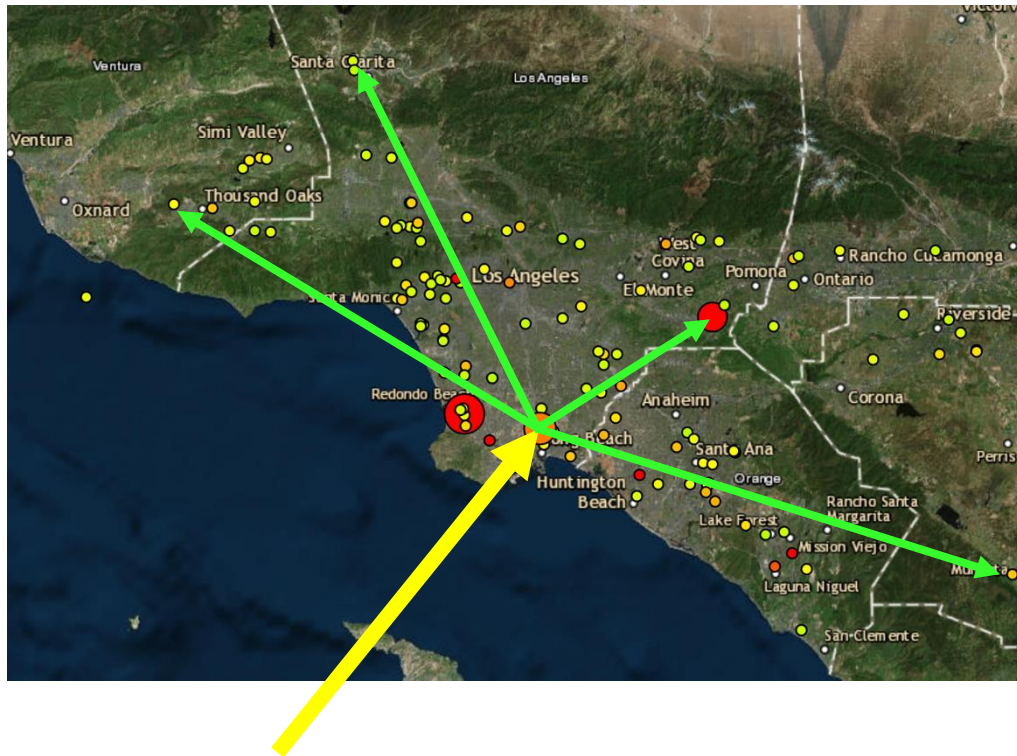
- Developed an insurance claim fraud model

Example: Spatial Data



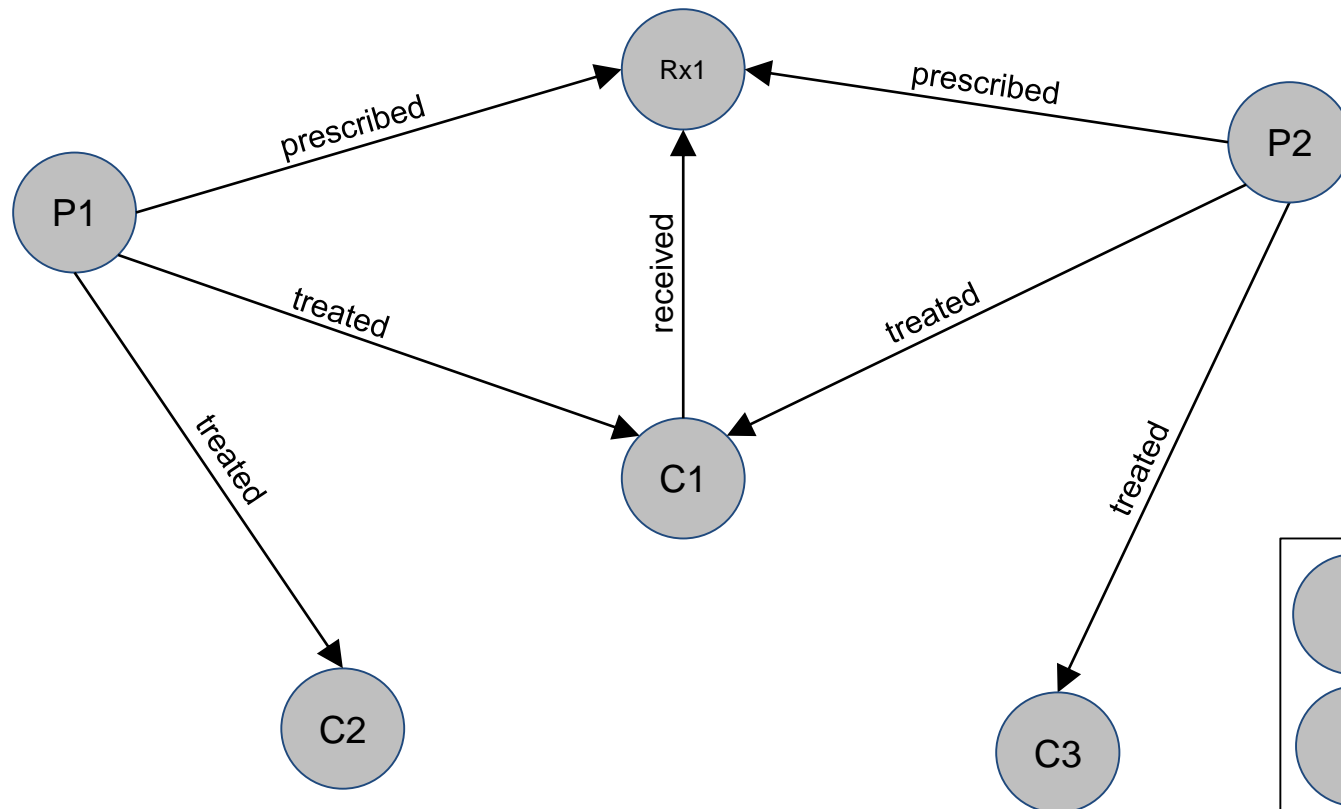
- Developed an insurance claim fraud model
- Distance traveled between *claimant* and *visited providers* became a feature

Example: Spatial Data

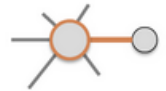


- Developed an insurance claim fraud model
- Distance traveled between *claimant* and *visited providers* became a feature
- Anomaly detection to look for unusual and unexpected behaviors

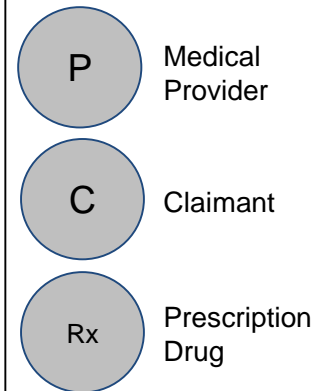
Example: Graph Data



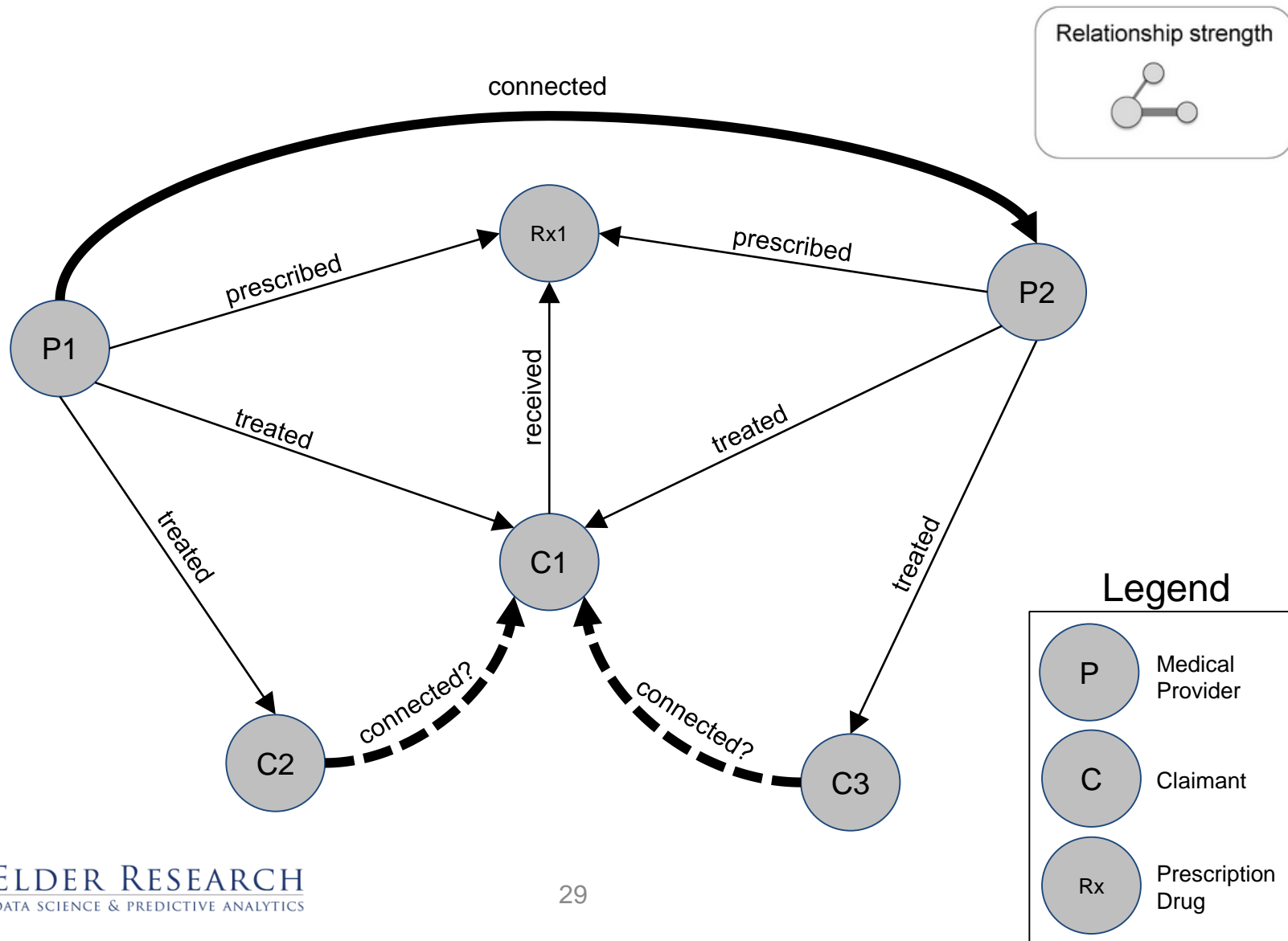
Key actors and links



Legend

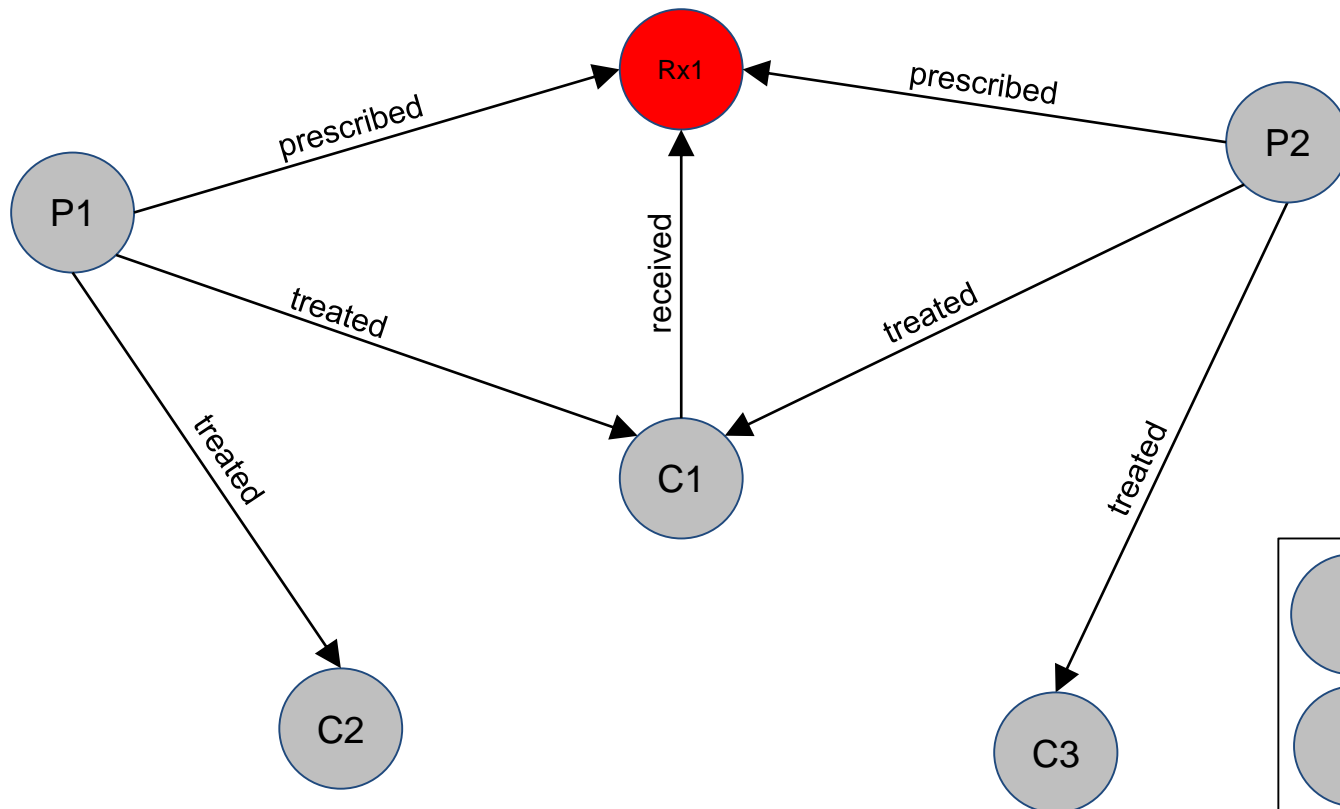


Example: Graph Data

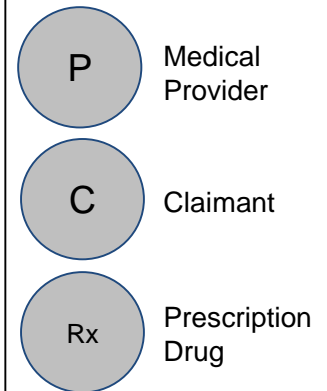


Example: Graph Data

Diffusion patterns

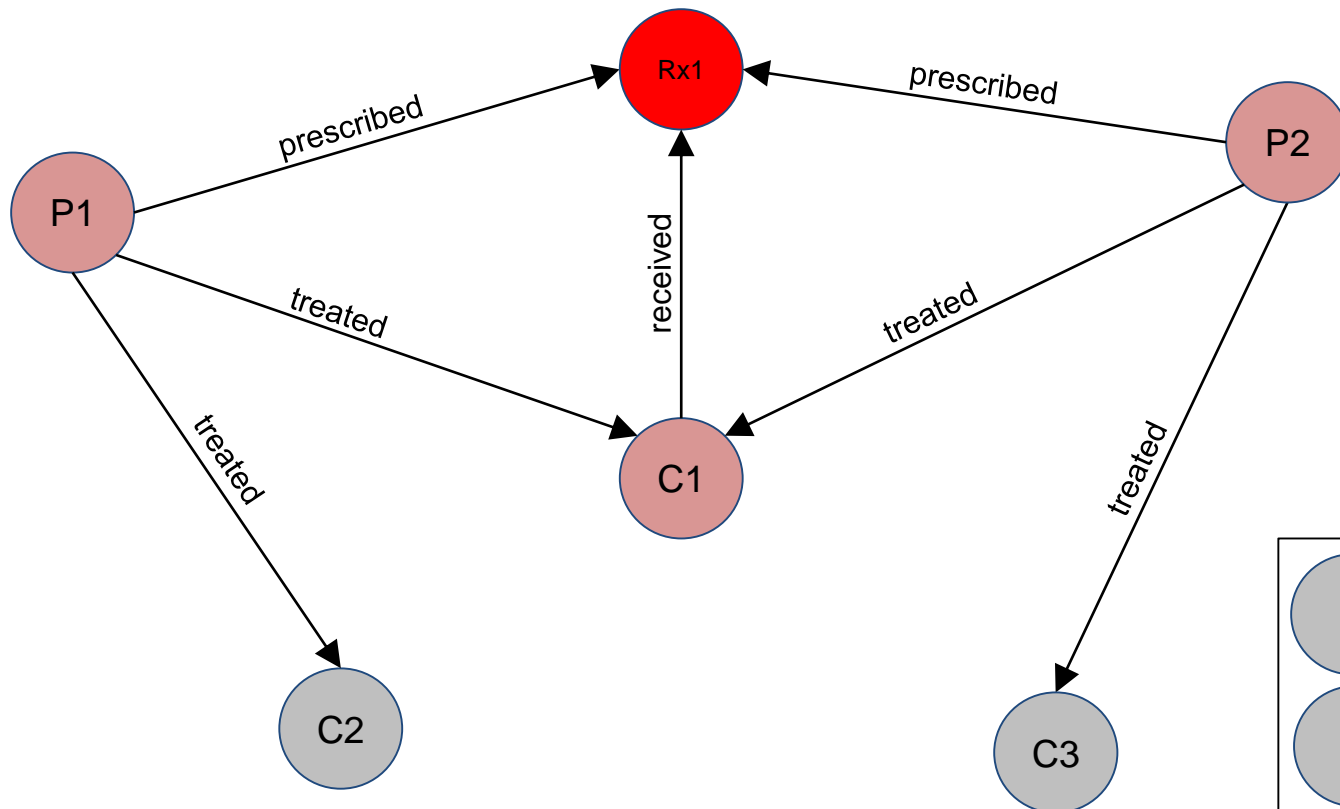


Legend

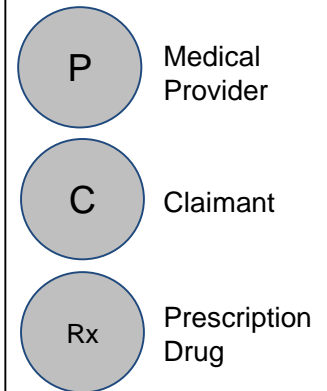


Example: Graph Data

Diffusion patterns

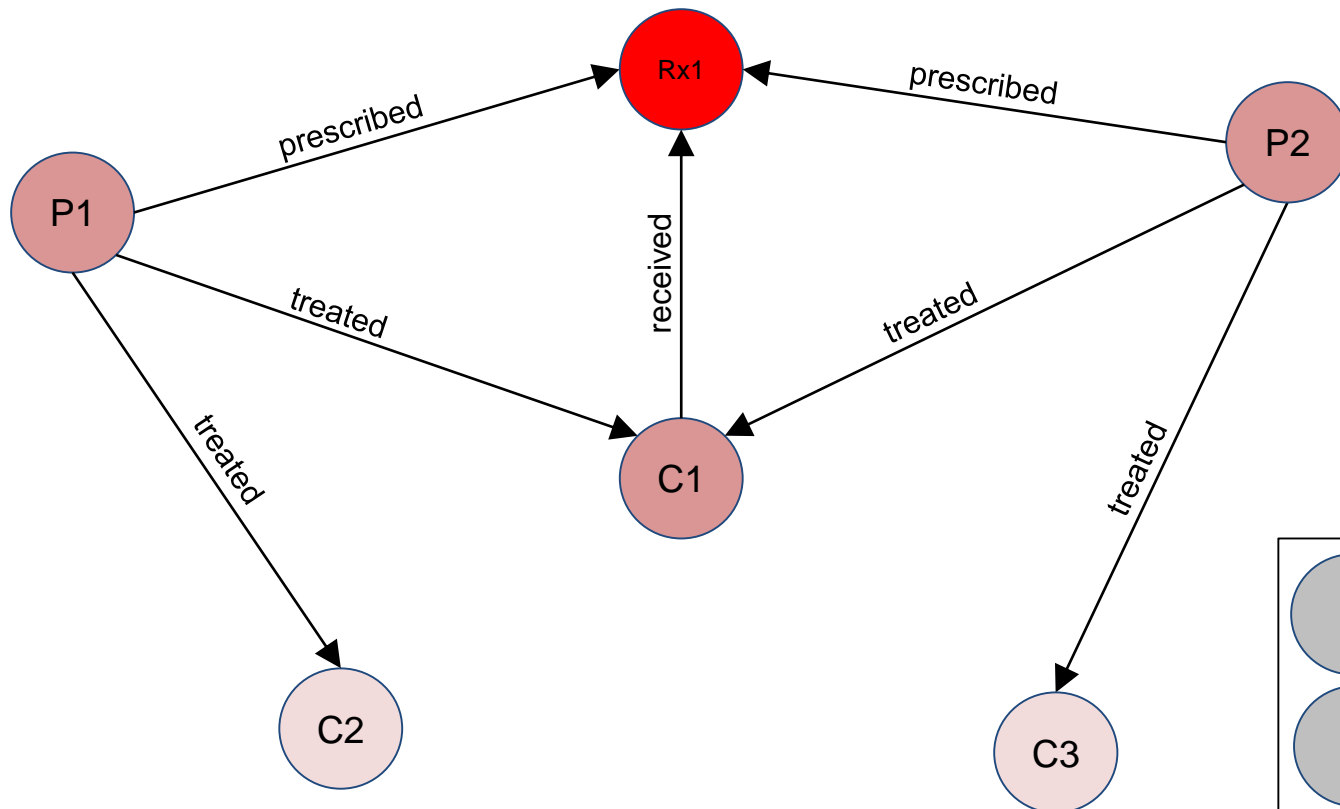


Legend

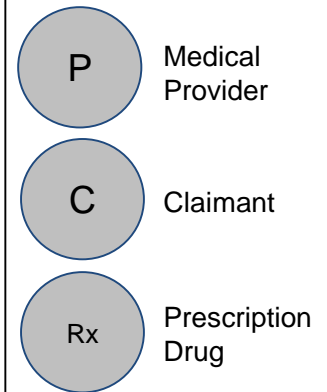


Example: Graph Data

Diffusion patterns



Legend



Move from model to solution

Consider Your Goals

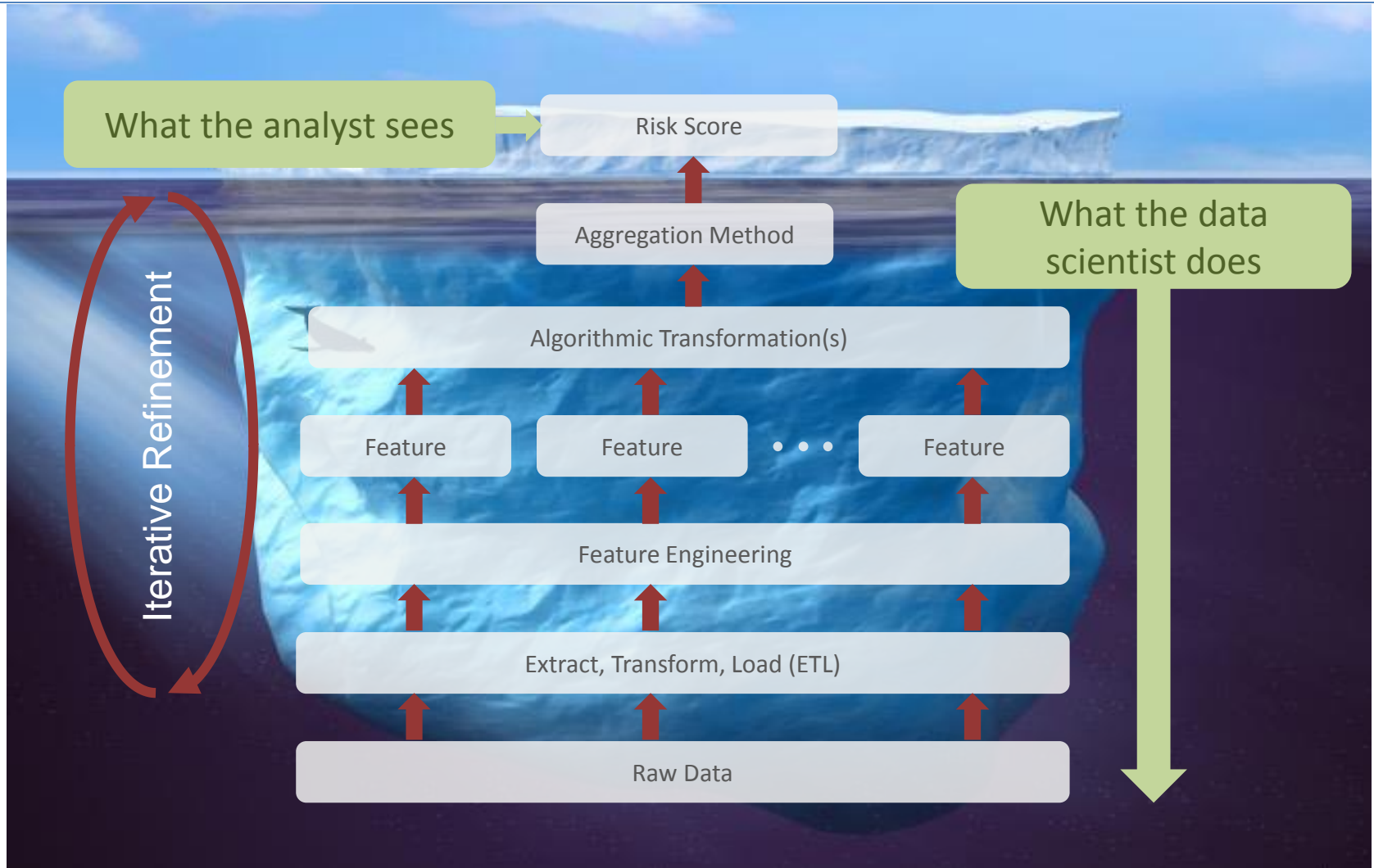
- What is your goal when deploying a fraud analytics solution?



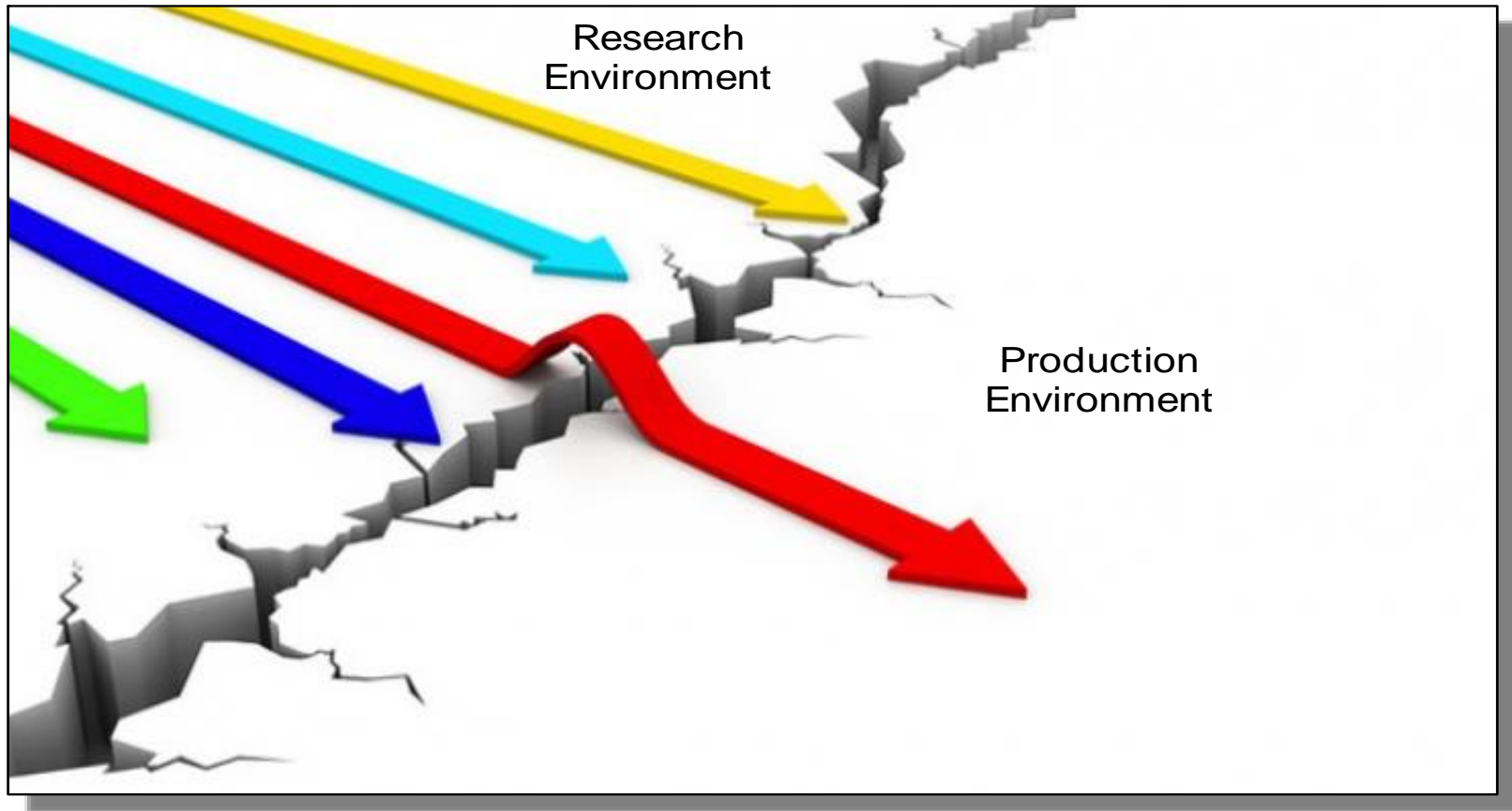
Move from model to solution

- “Model deployment” is not a machine learning problem, it’s a solution development problem
 - Who is the customer?
 - What is their job function?
 - What pain points do they experience regularly?
 - Are there efficiency gain opportunities?

Move from model to solution



Move from model to solution



Case Study

- Investigative Organization
- Excel spreadsheets with SSN's and risk scores (pretty good models)
- No traction and adoption with the customer (analysts and investigators)

Case Study

- Investigative Organization
- Excel spreadsheets with SSN's and risk scores (pretty good models)
- No traction and adoption with the customer (analysts and investigators)
- It didn't address a ***pain point*** – it created additional uncomfortable steps in their process

Case Study

- Took a customer-centric approach
 - Engaged the customer
 - Identified pain points
 - Built a solution around their natural workflow

Case Study

- Took a customer-centric approach
 - Engaged the customer
 - Identified pain points
 - Built a solution around their natural workflow
- Data access
- Data story telling
- “Simple” Automation

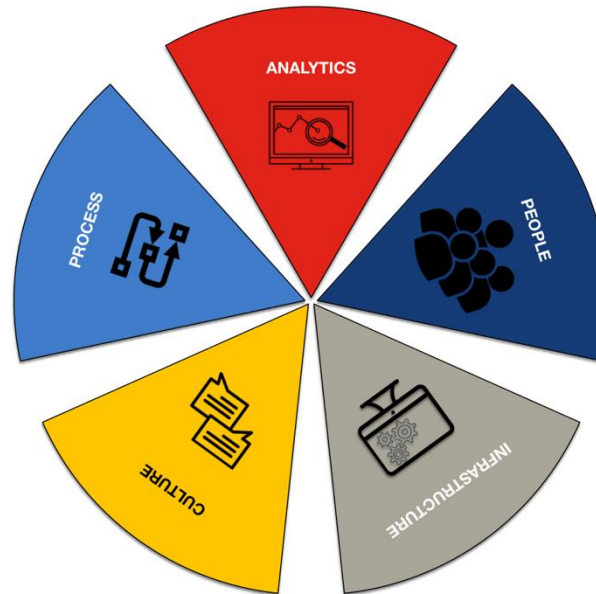
Case Study

- Took a customer-centric approach
 - Engaged the customer
 - Identified pain points
 - Built a solution around their natural workflow
- Data access
- Data story telling
- “Simple” Automation
- That led to **more** adoption of the solution and models

More than a good model...

More than a good model...

- Assess the maturity of your organization:



- Driving an entire organization to be more analytically minded and data-driven

Analytics as Disruptive Force

dis·rup·tive /dis'reptiv/

Adjective

Causing or tending to cause disruption

Synonyms

destructive



Case Study

- Two Federal Investigative Agencies
- We solved the same technical problem (very well) at both places
- However...
 - At Agency 1 → resulted in zero engagement
 - At Agency 2 → resulted in organizational adoption

Case Study

- Two Federal Investigative Agencies
- We solved the same technical problem (very well) at both places
- However...
 - At Agency 1 → resulted in zero engagement
 - At Agency 2 → resulted in organizational adoption



Case Study

- Infrastructure
 - IT environments, hardware, software, data access
- Process
 - Agile and CRISP-DM
- Analytics
 - Sophistication of modeling techniques utilized
- People
 - Equally talented and technically savvy customers

Case Study

- Culture.....well.....

Case Study

- Culture.....well.....
 - Agency 1
 - very political CYA leadership team
 - lacked organizational vision for analytics
 - no sense of mission urgency

Case Study

- Culture.....well.....
 - Agency 1
 - very political CYA leadership team
 - lacked organizational vision for analytics
 - no sense of mission urgency
 - Agency 2
 - cohesive leadership team
 - strong organizational vision for analytics
 - strong sense of urgency for the agency mission

Q & A

Upcoming Webinar

Detecting Fraud Rings with Graph Databases

Thursday, March 8th 2018 - 2:00-3:00 PM (EST)

This webinar will focus on how to identify suspicious behavior using tools and technology suited for network analysis, including graph databases.



Webinar hosts Robert Han (Director and Program Manager, Washington DC) and Ryan McGibony (Senior Data Scientist) have extensive experience working on fraud analytics projects across the federal government and commercial businesses.



Robert Han

Elder Research

Director and Program Manager

han@elderresearch.com



Miriam Friedel, Ph.D.

Elder Research

Director & Senior Scientist

miriam.friedel@elderresearch.com



Bryan Jones

Strategy First Analytics

Owner, Principal Consultant

bryan@strategyfirstanalytics.com

Visit our blog at www.elderresearch.com/company/blog