RSAConference2015
San Francisco | April 20-24 | Moscone Center

SESSION ID: ECO-T08

Majority Report: Making Security Data Actionable (and Fun!)

Andrew Hay
Research Director
OpenDNS, Inc.
@andrewsmhay

Thibault Reuille
Sr. Security Researcher
OpenDNS, Inc.
@thibaultreuille

RSAConference2015

San Francisco | April 20-24 | Moscone Center

#RSAC

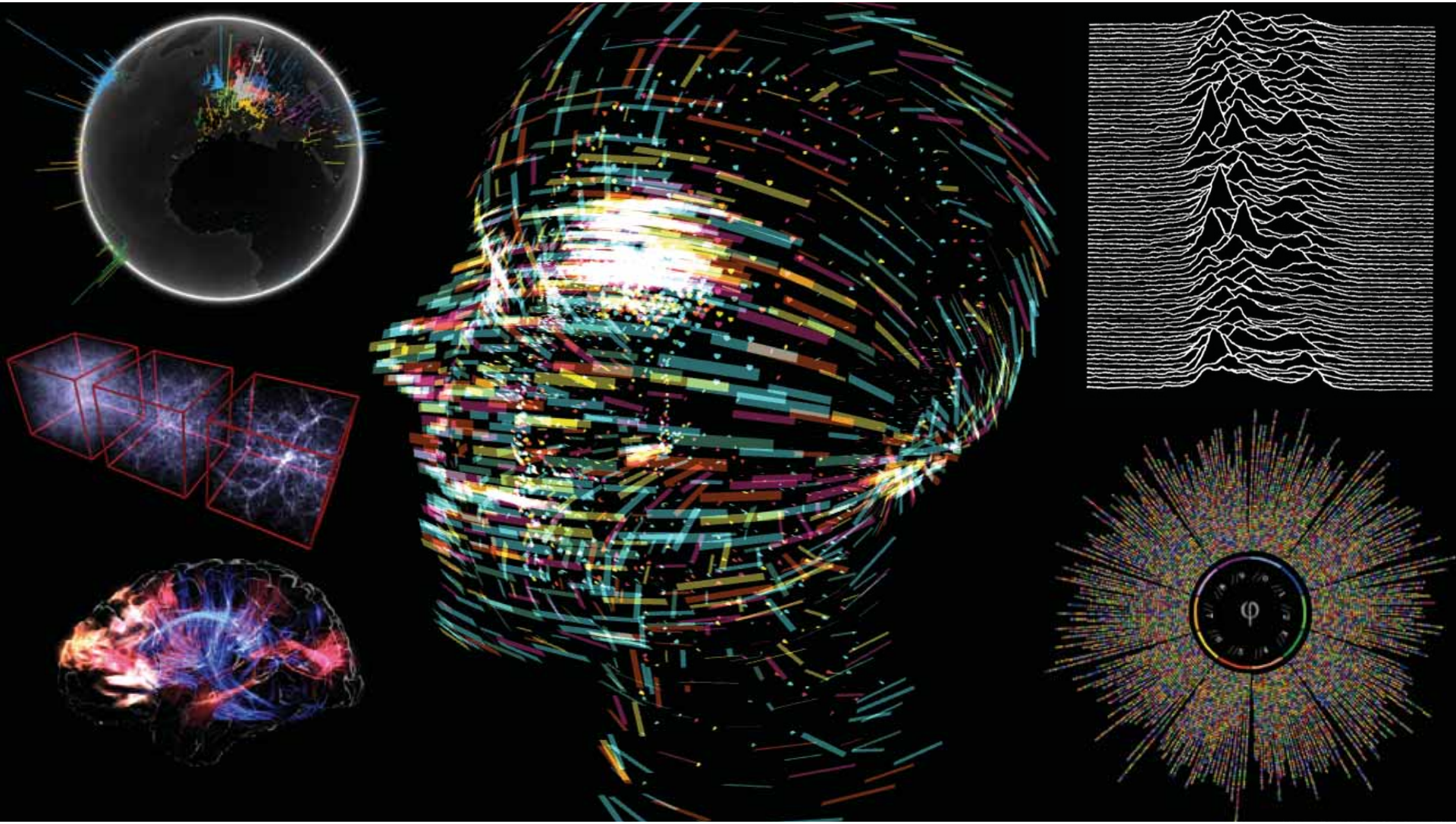
Why Visualize Data?

The Age of Information

Why Visualize the Data?

- Aren't pie charts enough?
- What does advanced visualization give us?
- Can't I just use R, Python, and/or Excel?

Because, Minority Report



Is Data the New Oil?

- Collection
- Storage
- Persistence

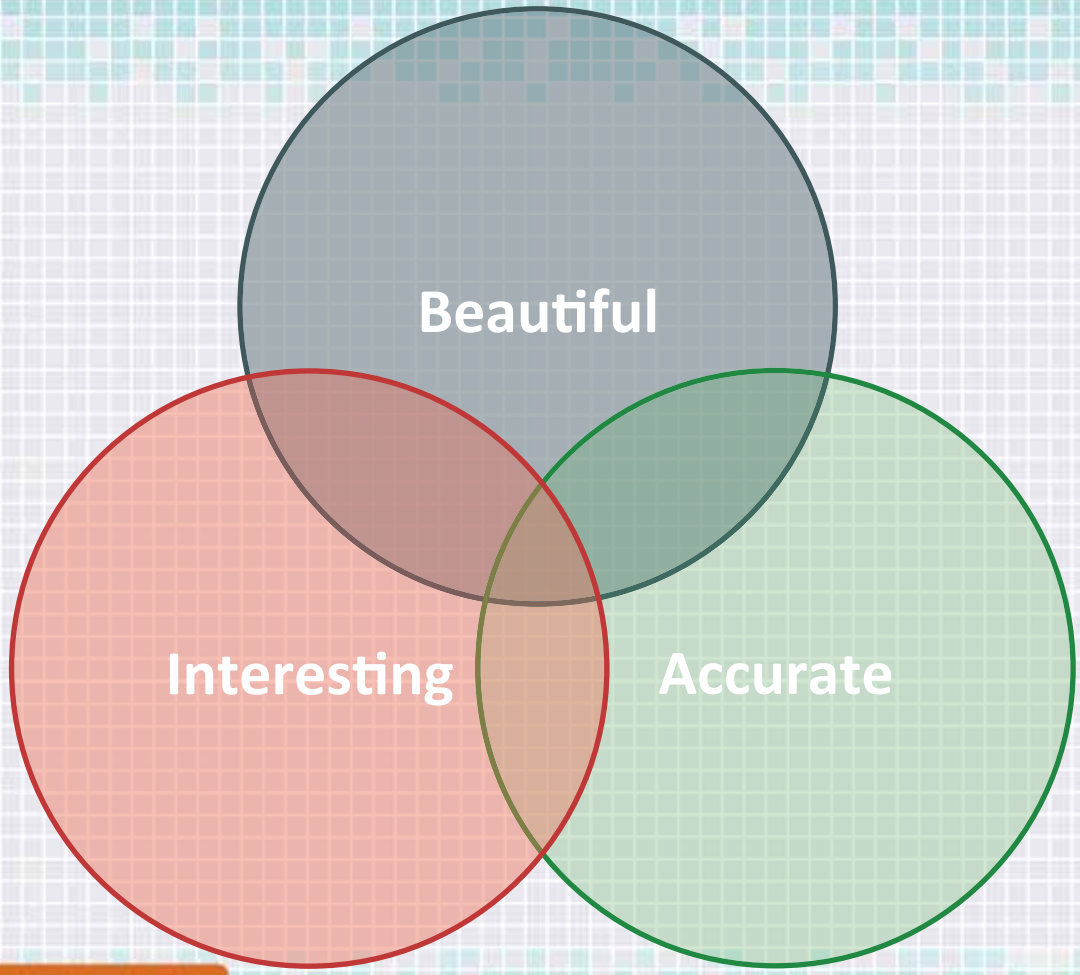
- Analytics
- Statistics
- Machine Learning

- Indicators
- Insights
- Stories

"There are 3 kinds of lies: Lies, damned lies, and statistics."

- Benjamin Disraeli, 19th Century British Prime Minster

Image source: http://en.wikipedia.org/wiki/Benjamin_Disraeli#/media/File:Benjamin_Disraeli_by_Cornelius_Jabez_Hughes,_1878.jpg



- Crystal Clear Design
- Relevant & Useful
- Scientific Method

RSAConference2015

San Francisco | April 20-24 | Moscone Center

#RSAC

Quick Overview of Learning Styles

Learning Styles

- Neil Fleming's VAK/VARK model

- The 4 types
  1. Visual learners
  2. Auditory learners
  3. Reading-writing preference learners
  4. Kinesthetic learners or tactile learners

Learning Styles

- Key concept of visual learning
- Graphic organizers
- Visual representations of
  - knowledge,
  - concepts,
  - thoughts, or
  - ideas

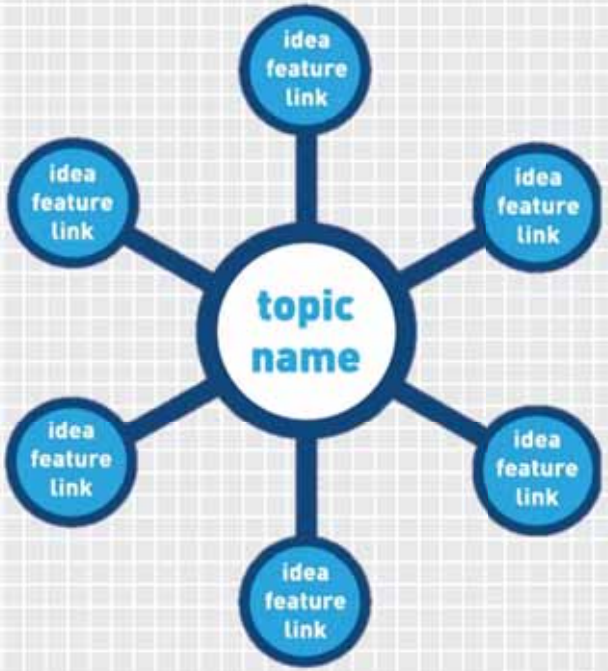
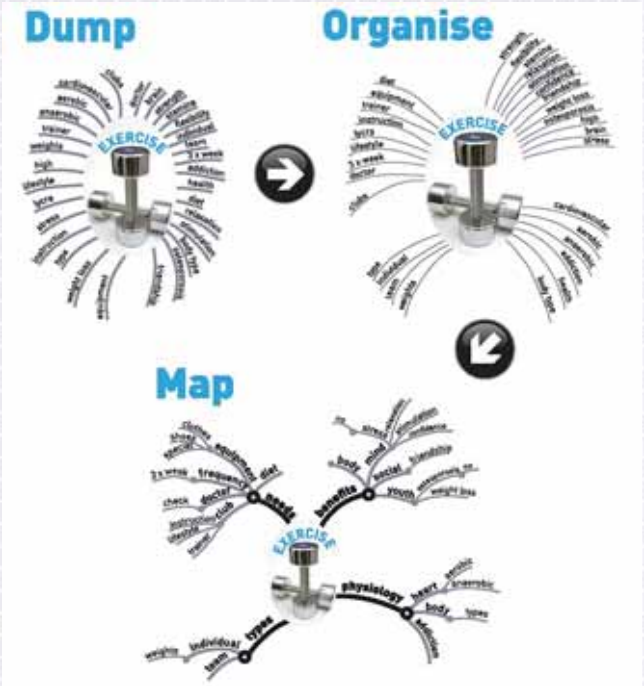


Photo Credit: modellearning

Learning Styles



- Clarify meaning through relationships
- Best example might be utilizing a mind map

Photo Credit: modellearning

Learning Styles

- Representing information spatially and with images [some*] students are able to
  - focus on meaning
  - reorganize and group similar ideas easily
  - make better use of their visual memory

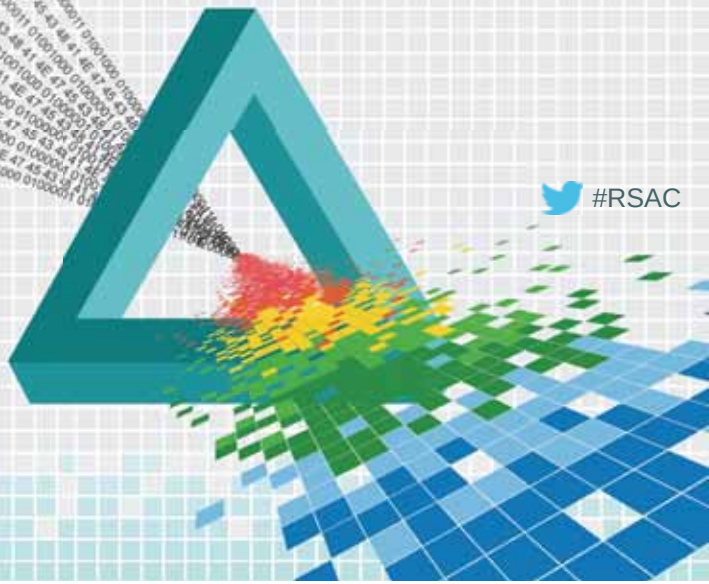
Source: http://en.wikipedia.org/wiki/Visual_learning

RSAConference2015

San Francisco | April 20-24 | Moscone Center

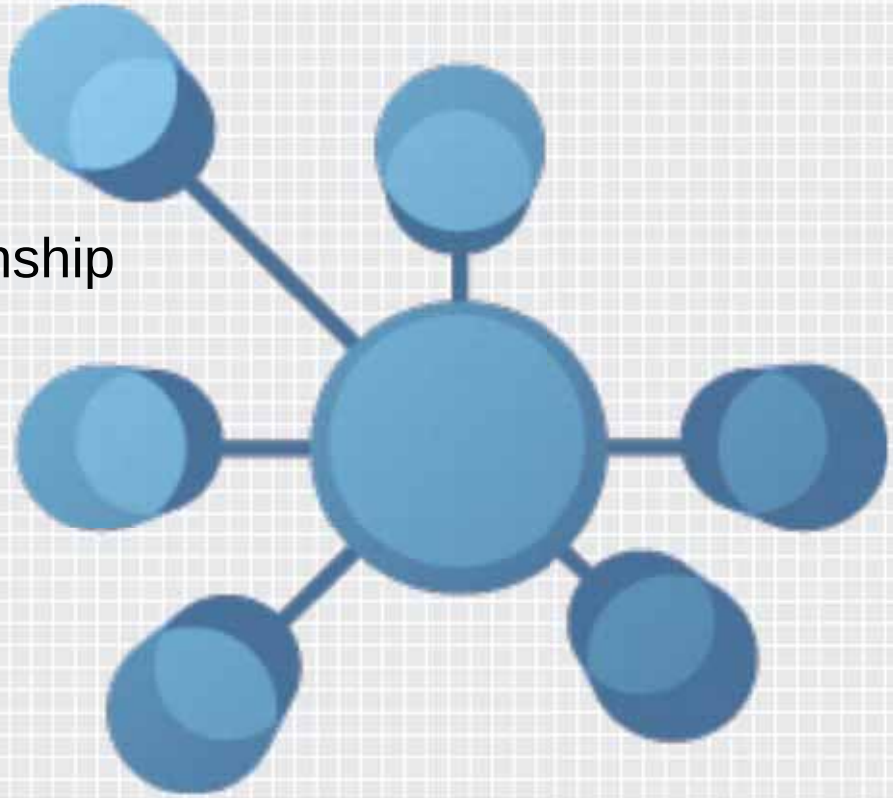
#RSAC

Visualization Fundamentals

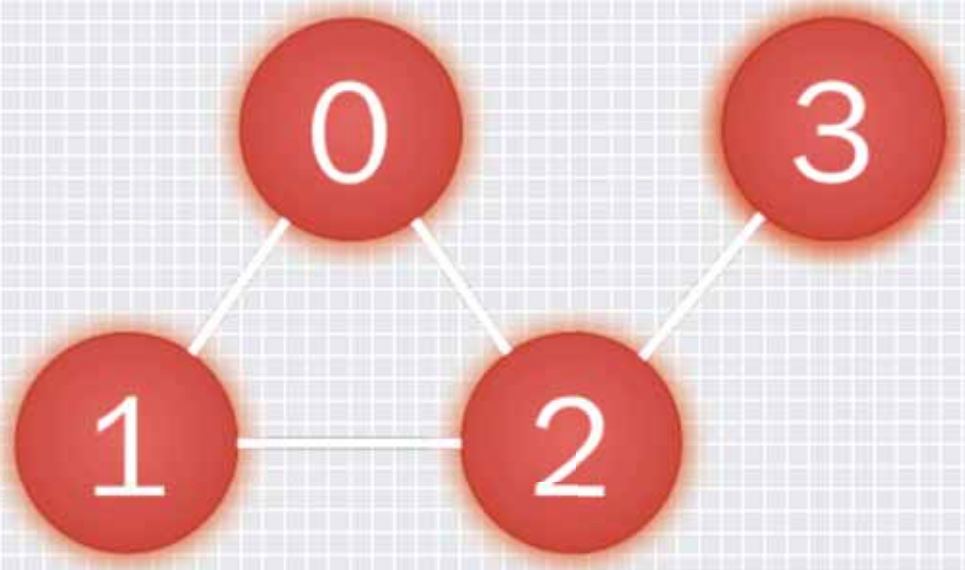


Knowledge

- Semantic Networks
- Node = Concept, Edge = Relationship
- Model of the Information
- Ontology
  - Model of the Model



Graph Theory 101



Let G be a graph

G = (V, E)

Where

V = {0, 1, 2, 3}

And

E = {(0, 1), (0, 2), (1, 2), (2, 3)}

Graphs Are Everywhere!

- Social Networks
- Maps & Shortest Path
- UML Software Design
- Neural Networks

RSA Conference2015

San Francisco | April 20-24 | Moscone Center

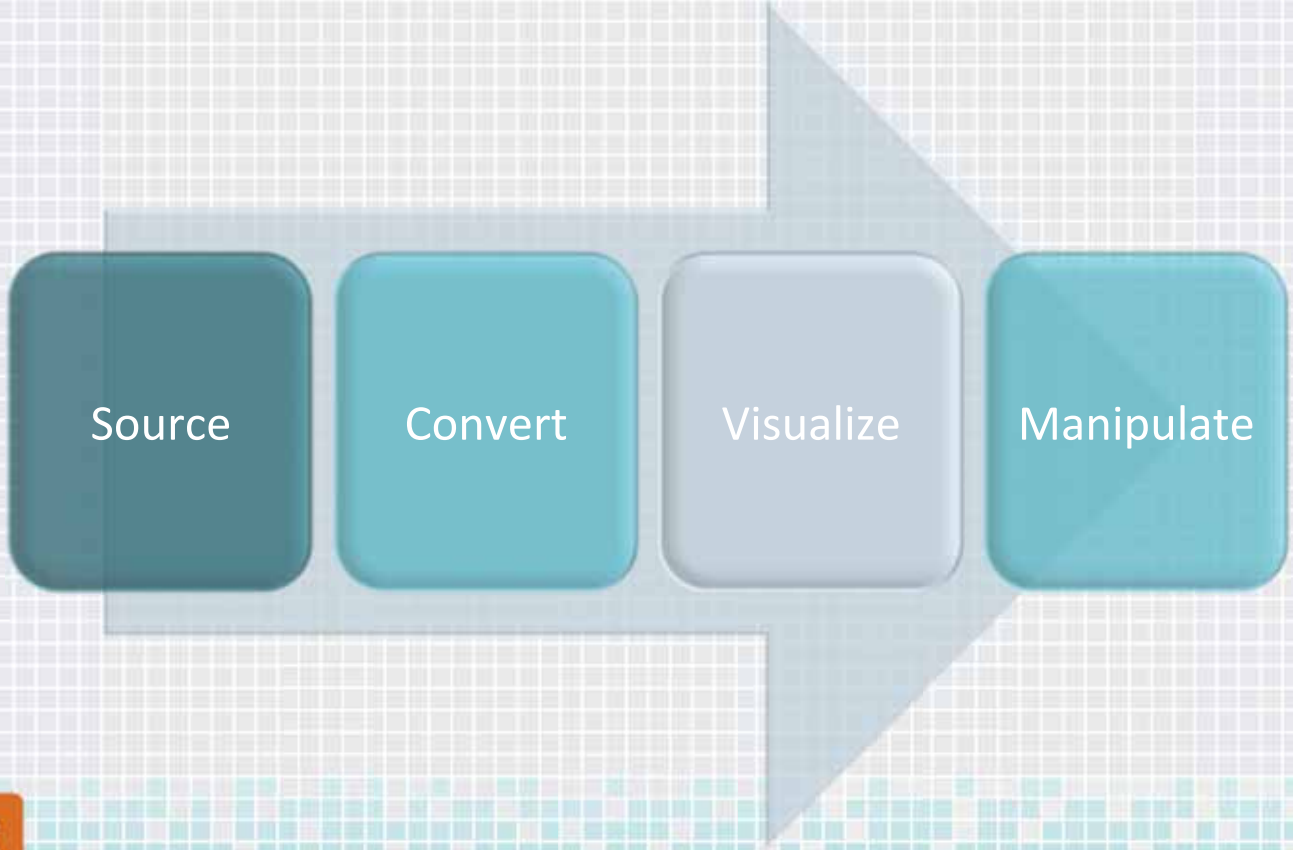
Introducing OpenGraphiti

#RSAC

Introducing OpenGraphiti

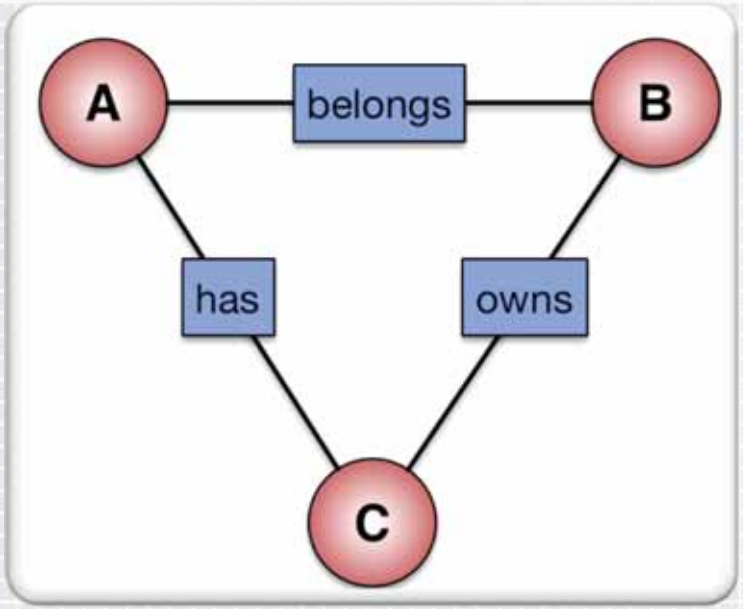
- 3D data visualization engine
- Free & Open Source
- Visualize and manipulate any loosely related data
- Easy-to-use shell for data scientists

Workflow

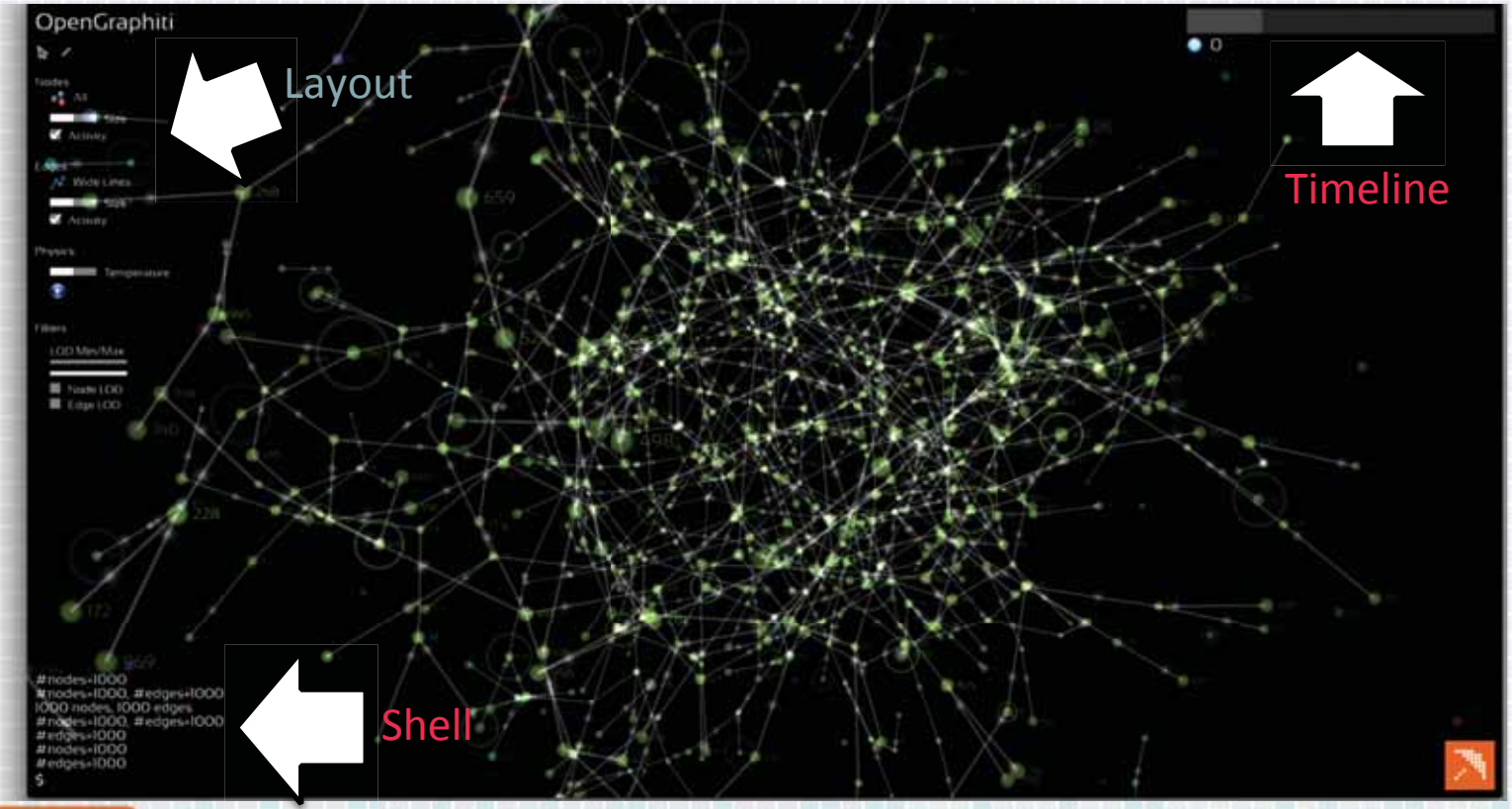


SemanticNet Library

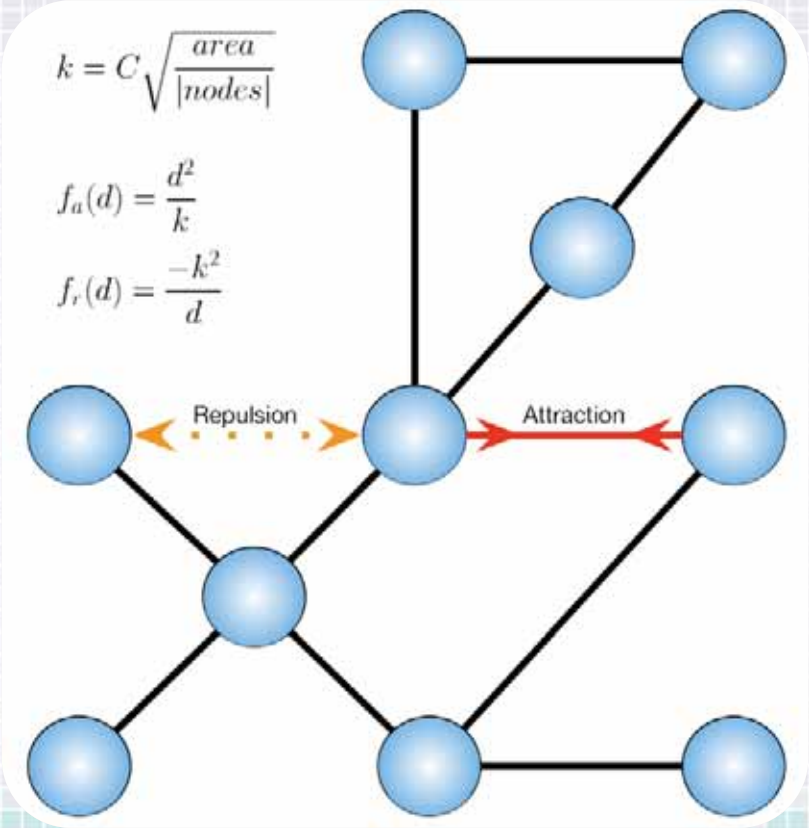
```
#!/usr/bin/env python

import semanticnet as sn

graph = sn.Graph()

a = graph.add_node({ "label" : “A” })
b = graph.add_node({ "label" : “B” })
c = graph.add_node({ "label" : “C” })

graph.add_edge(a, b, { "type" : “belongs” })
graph.add_edge(b, c, { "type" : “owns” })
graph.add_edge(c, a, { "type" : “has”

graph.save_json(“dataset.json”)
```

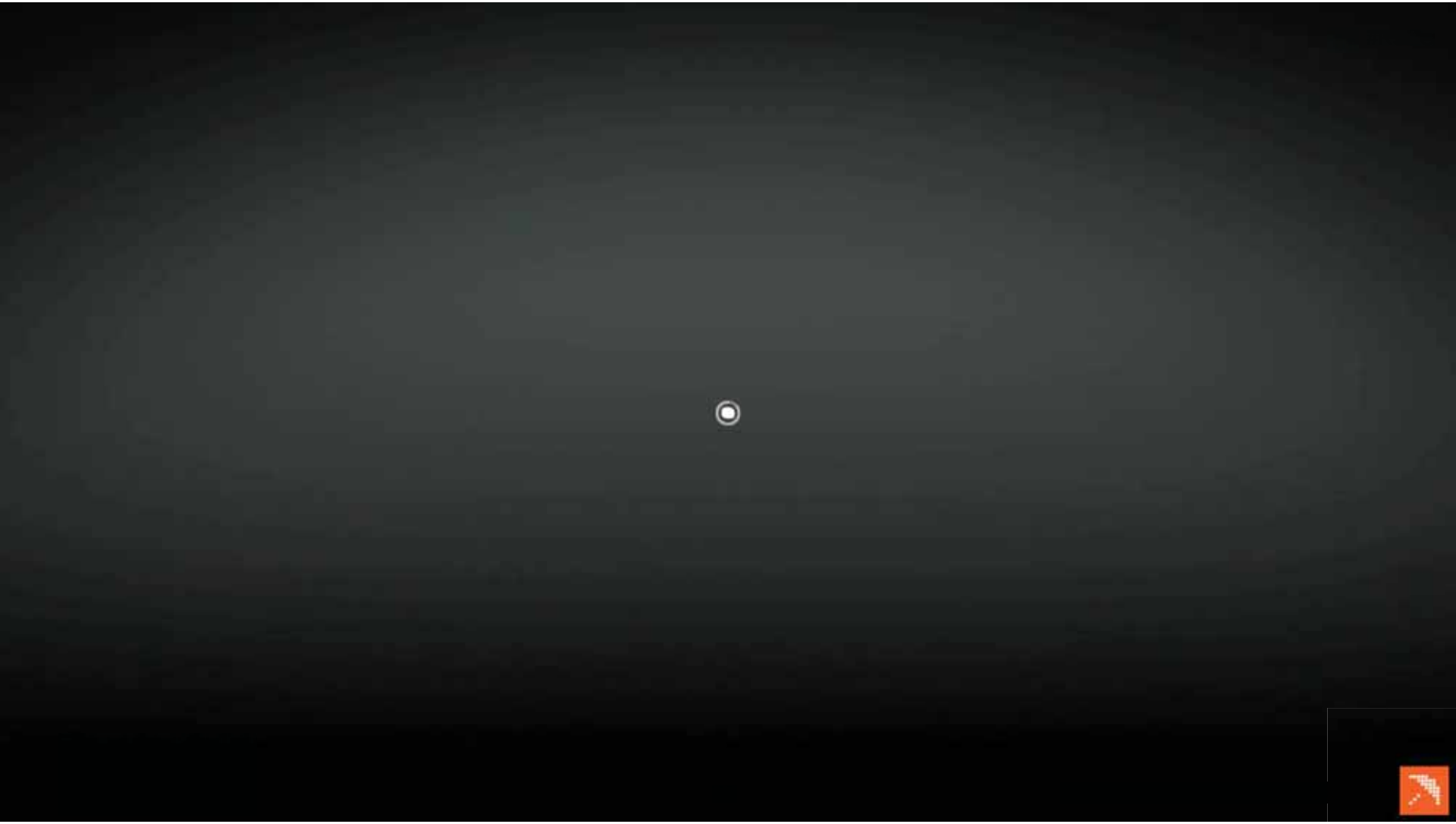


OpenGraphiti



Particle Physics



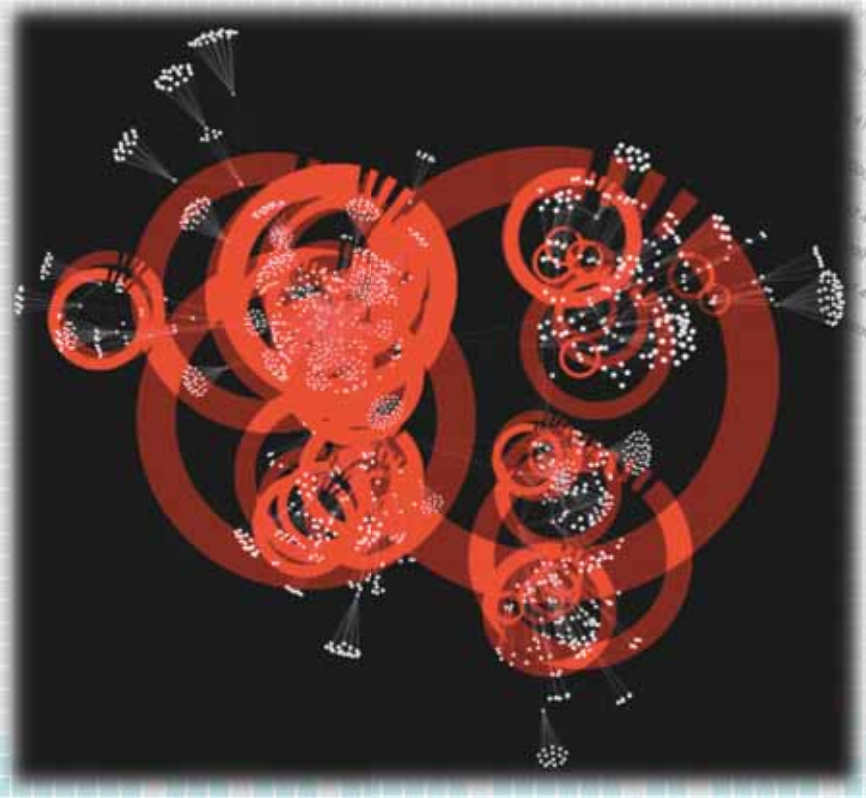


Why?

- Actors populate the knowledge graph
- Creation is understood, output is complex
- Layout closer to the *“natural shape”* of data structure
- Take advantage of the GPU to untangle information
- Humans are good at processing shapes and colors

RSA Conference2015

San Francisco | April 20-24 | Moscone Center

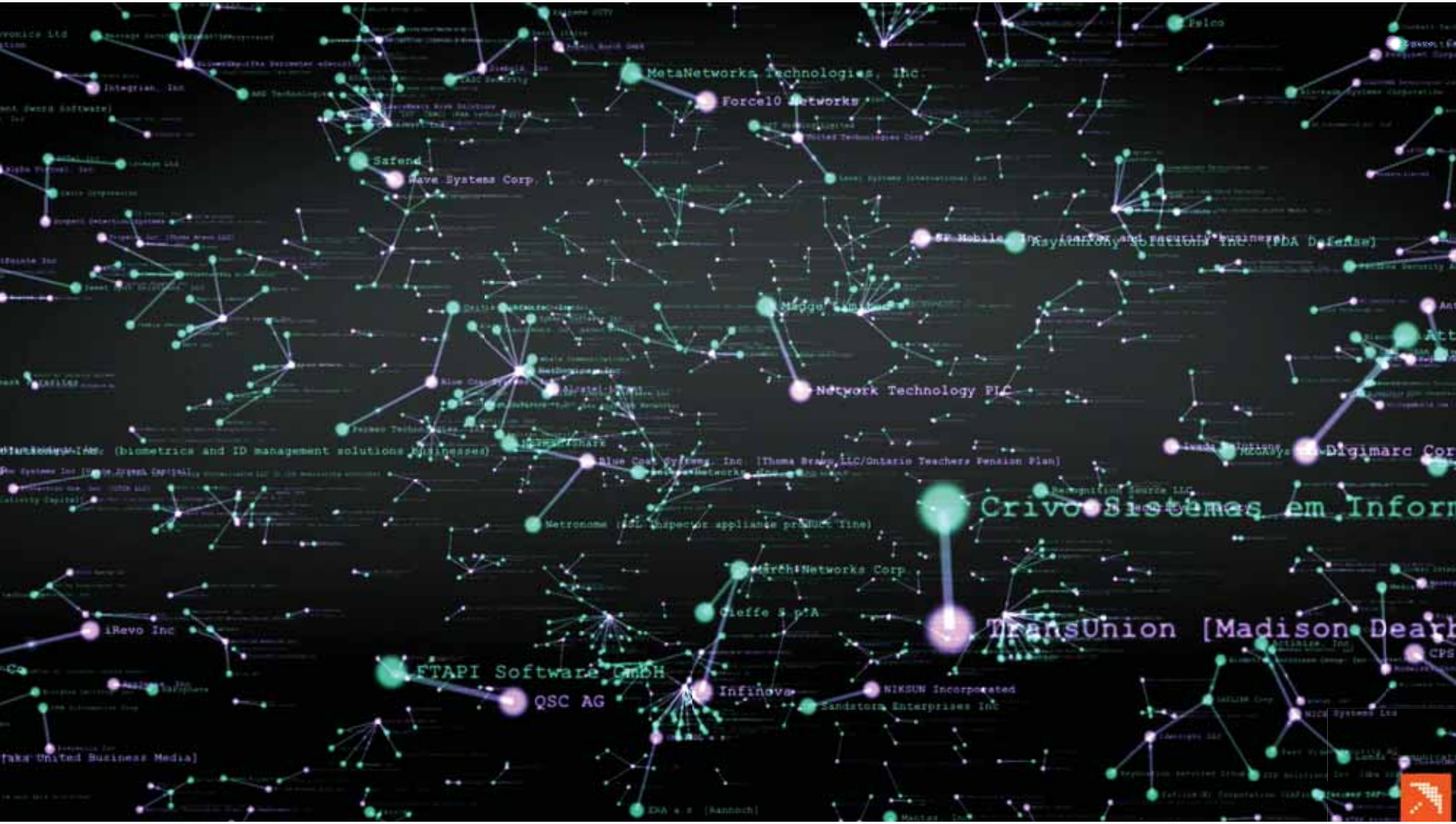


Datasets!

#RSAC

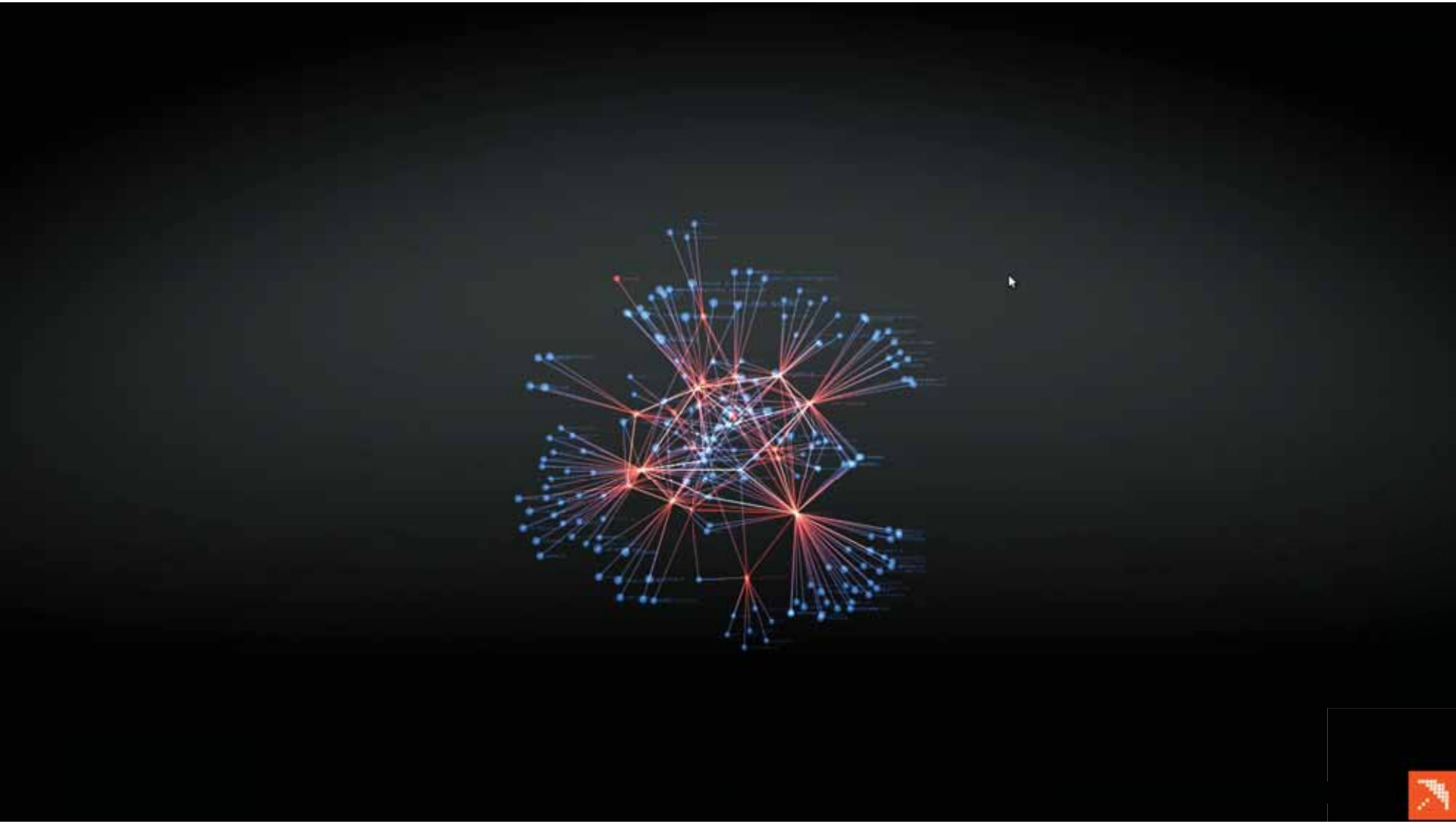
Merger & Acquisition Dataset

- Looked at all merger and acquisition activity in the security space since 2002 (*to the summer of 2014*)
- Connected:
  - Acquirers (who did the acquisition)
  - Targets (who was acquired)
  - Acquisition value (if disclosed)
- Data sourced from 451 Research's M&A Knowledgebase
  - https://451research.com/merger-aquisitions-knowledgebase-overview


MetaNetworks Technologies, Inc.
Force10 Networks
Safend
Network Technology PLC
Crivo Sistemas em Inform
TransUnion [Madison Dearb
FTAPI Software GmbH
QSC AG
Infinova
NIKSUN Incorporated
Sandstorm Enterprises Inc
iRevo Inc

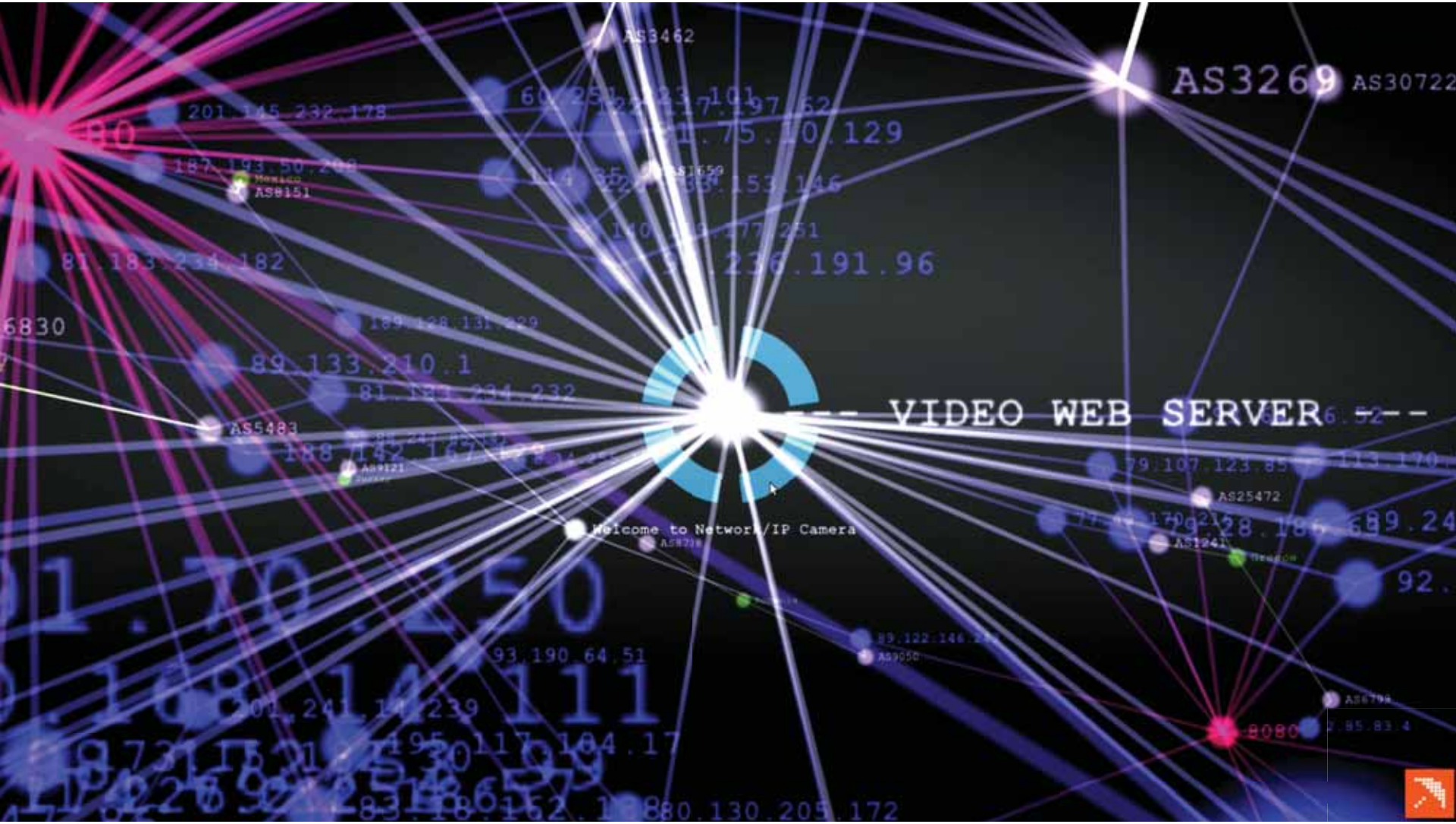
Visualizing a Partner Network

- Graph of vendors and their partners
- Looking at a security partner ecosystem
- Who should *your* company partner with?



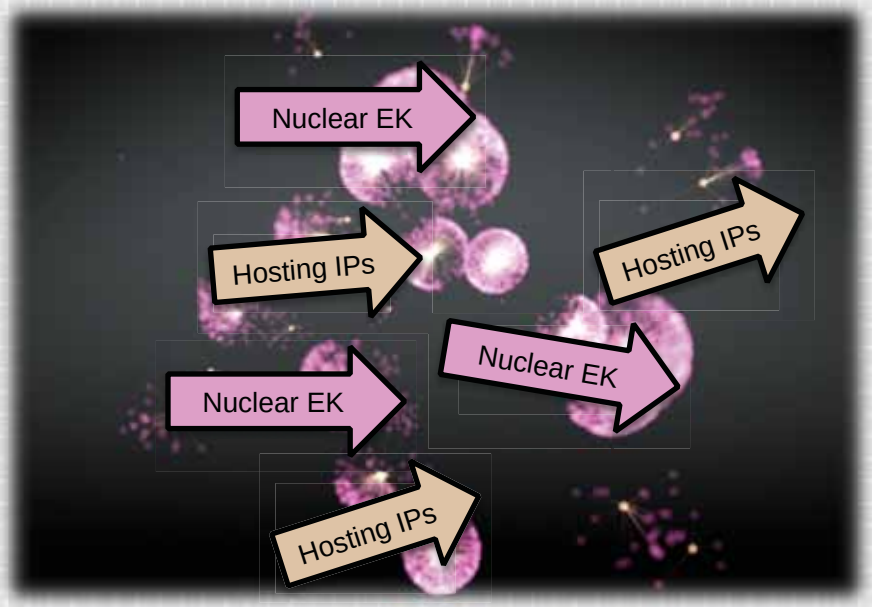
Visualizing a SHODAN Query

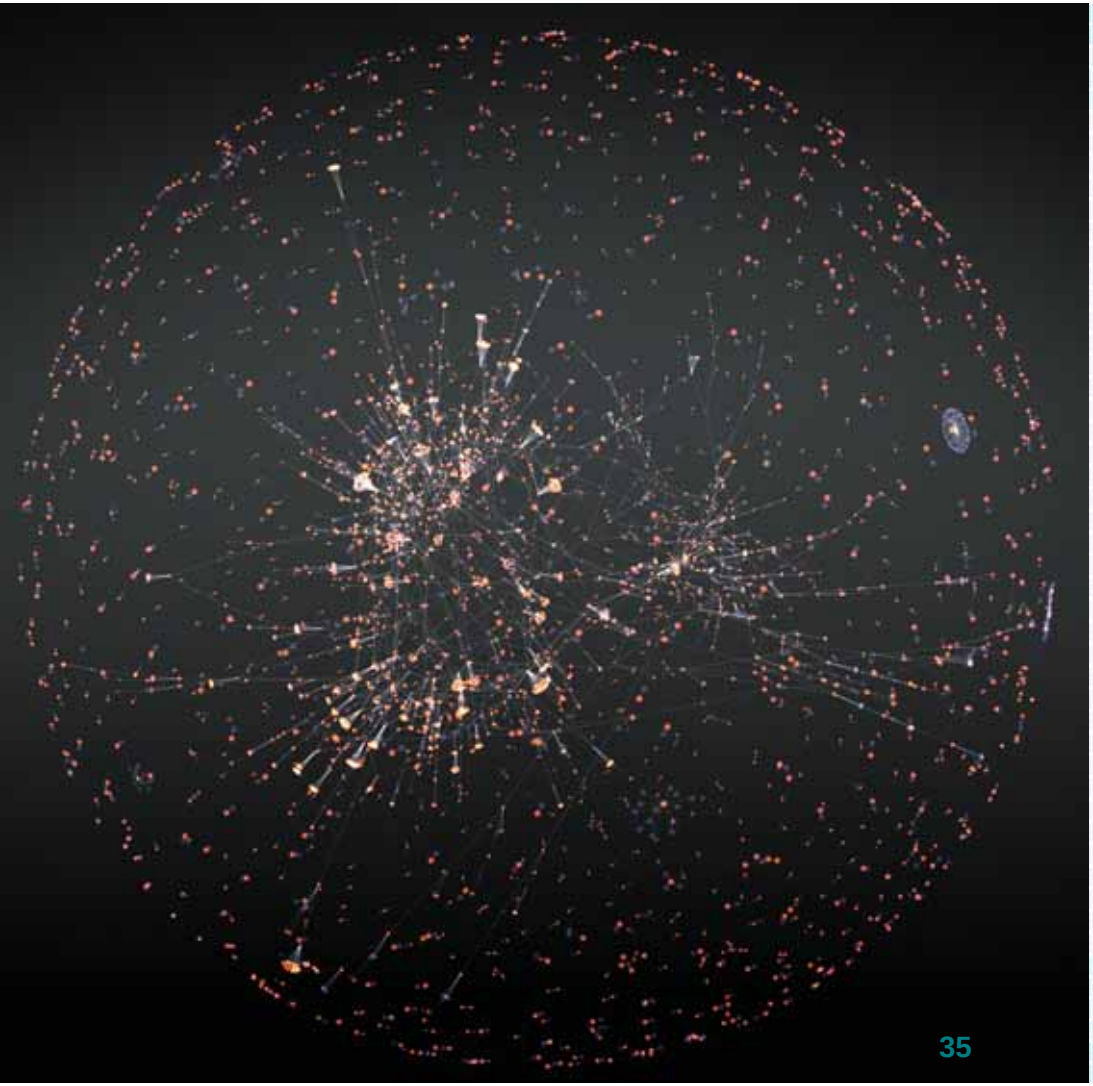
- https://www.shodan.io/
- Lets you find specific computers (routers, servers, etc.) using a variety of filters
- Some have described it as a public port scan directory or a search engine of banners


AS3462
AS3269
AS30722
AS8151
AS5483
AS25472
VIDEO WEB SERVER
Welcome to Networ /IP Camera
8080

Nuclear Exploit Kit Dataset

- List of domains hosting the Nuclear EK (pink)
- Retrieve their IPs (yellow)
- Create a Domain-IP graph
- Quickly see the clusters





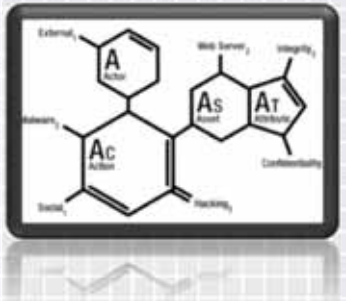
Email Graph

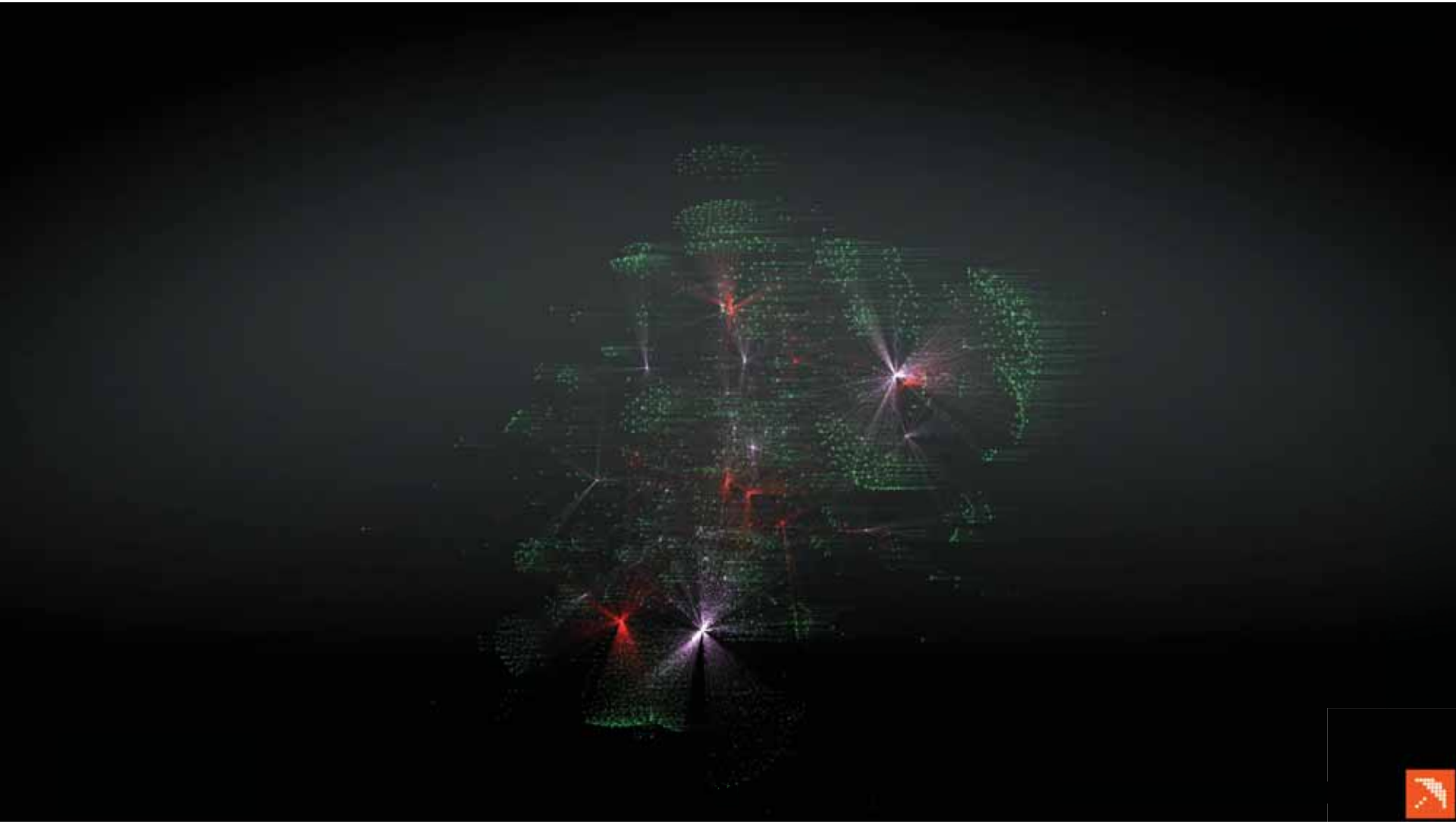
- Graph of email communications
- Indicative of distinct
  - Silos/divisions
    - Different locations?
  - Many-to-one comms
    - Email forwarding to team?
  - Orphan nodes
    - Spam?

Email Graph

VCDB Dataset

- http://www.vcdb.org
- From the Verizon Risk Team
- Vocabulary for Event Recording and Incident Sharing (VERIS)
- VERIS Community Database (VCDB)

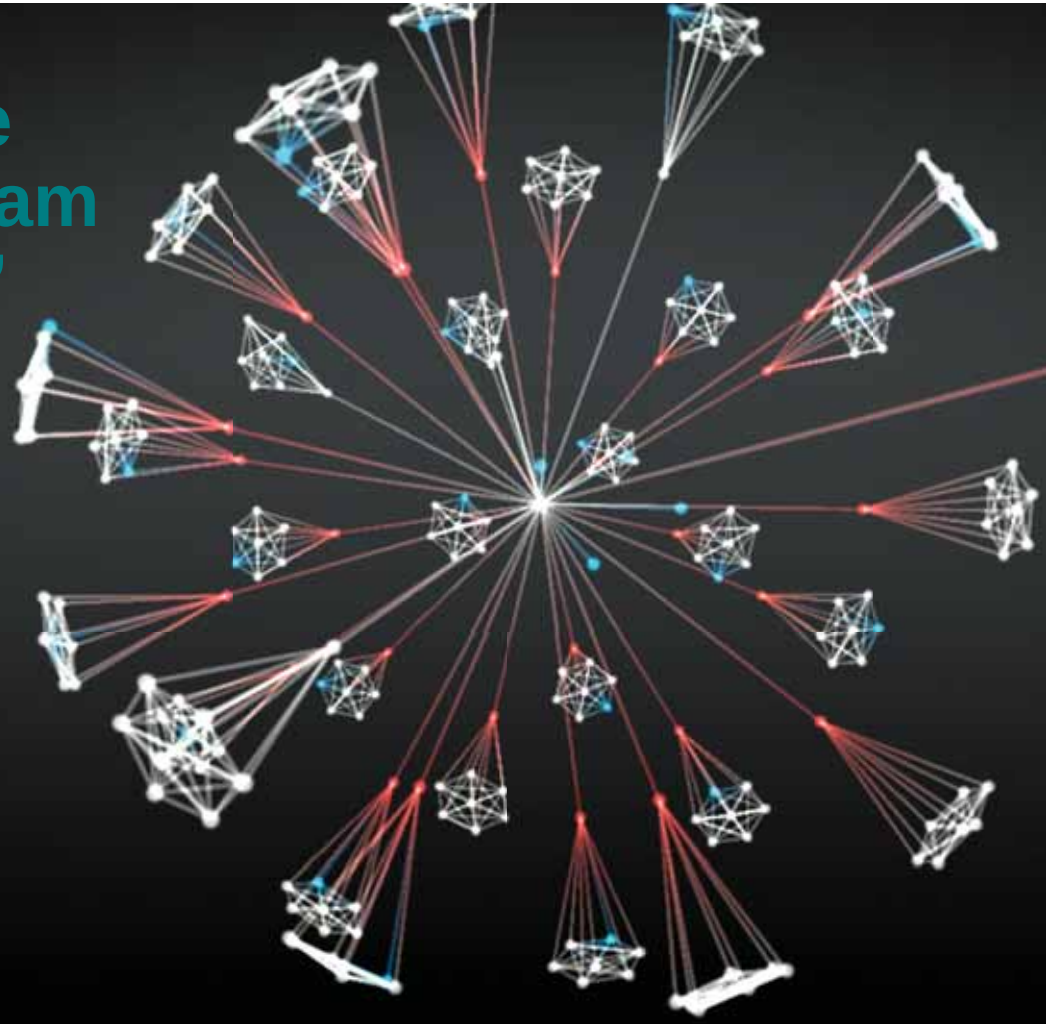




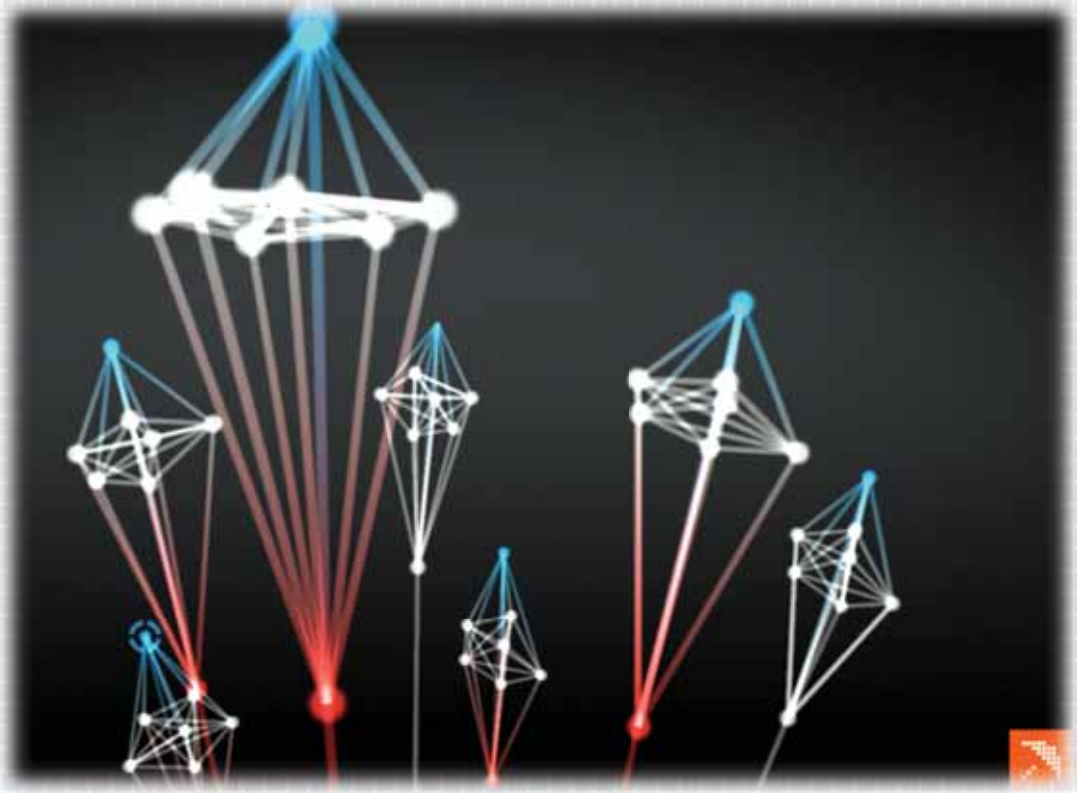
Dyre Botnet

- Typically distributed by the Cutwail botnet
  - Via spam emails
  - With links to Dropbox or Cubby files
- Harvests credentials through traffic interception
  - Primarily targets online banking websites for ACH/wire fraud
- When graphed, a we noticed the emergence of a unique shape

The Dyre ‘Pentagram Network’



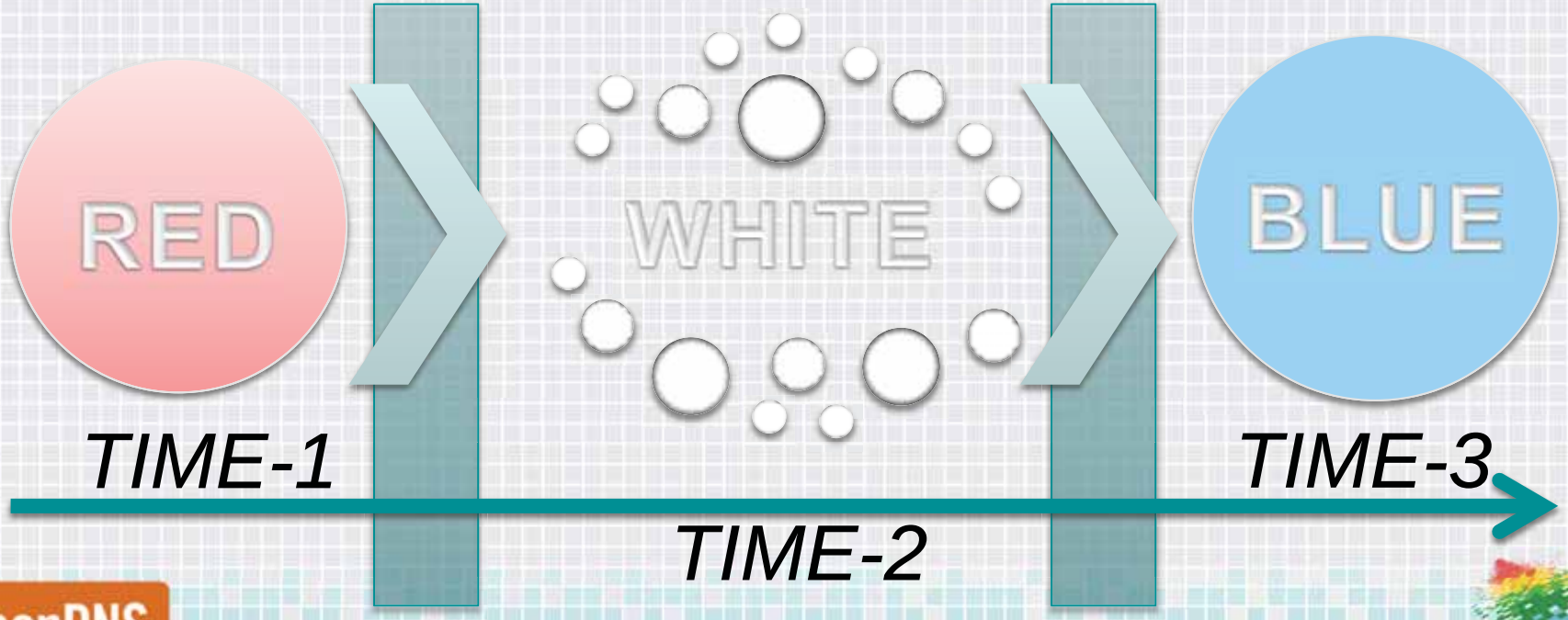
The Dyre ‘Pentagram Network’



- Initial query to domain (red)
- Co-occurrences with domains in a ‘pentagram network’ (white)
- Finally directed out of ‘pentagram network’ (blue)

OpenDNS

Linear Queries & Time Compression



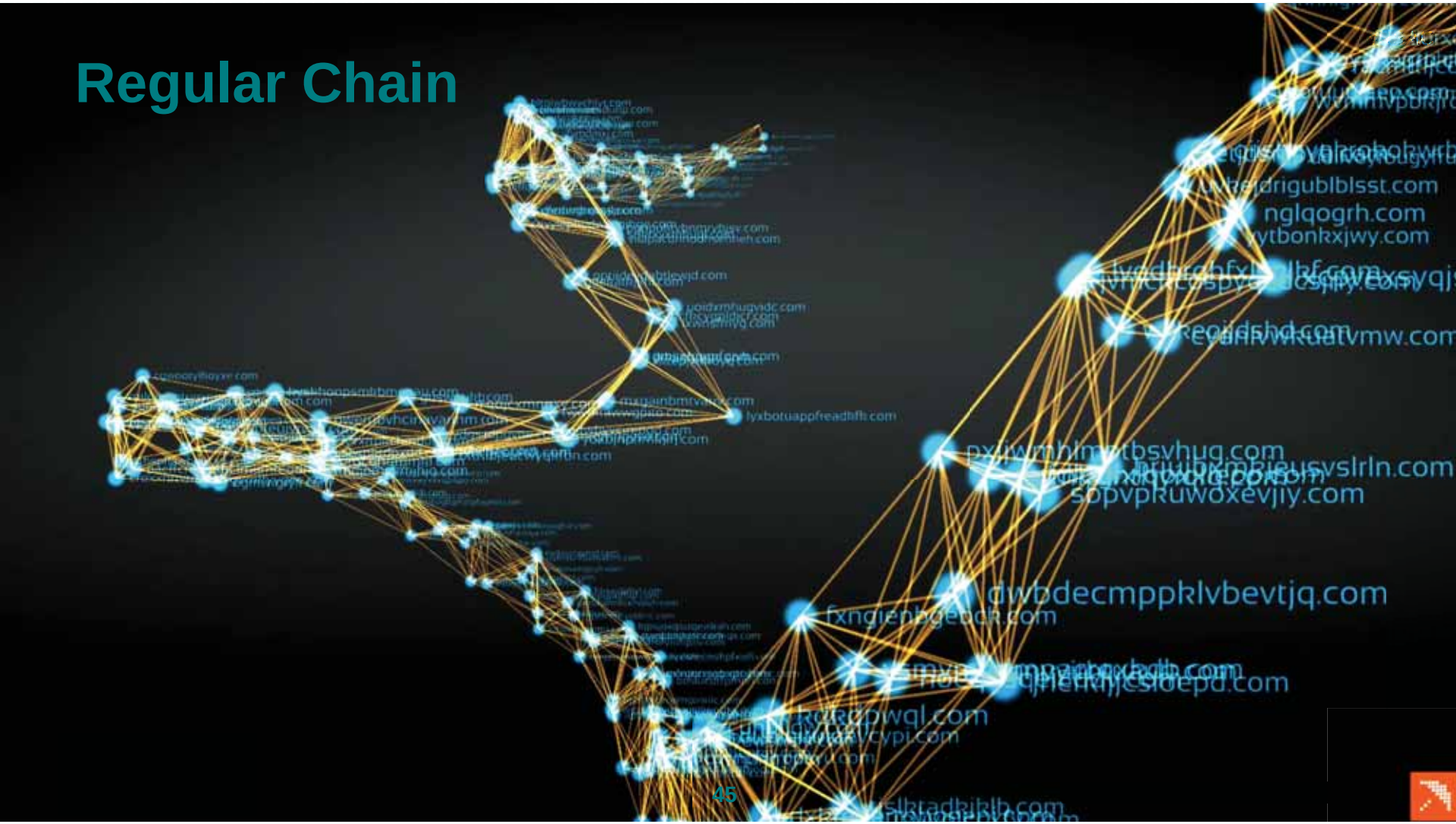
Exploring DGAs

- Domain Generation Algorithm (DGA)
- Used to generate domains programmatically
- Typically rely on a seed of some sort
  - Date, time, keyword, etc.
  - Allows for the registration of domains that no human would ever type
- Not just used for malware
  - Also used for marketing campaigns

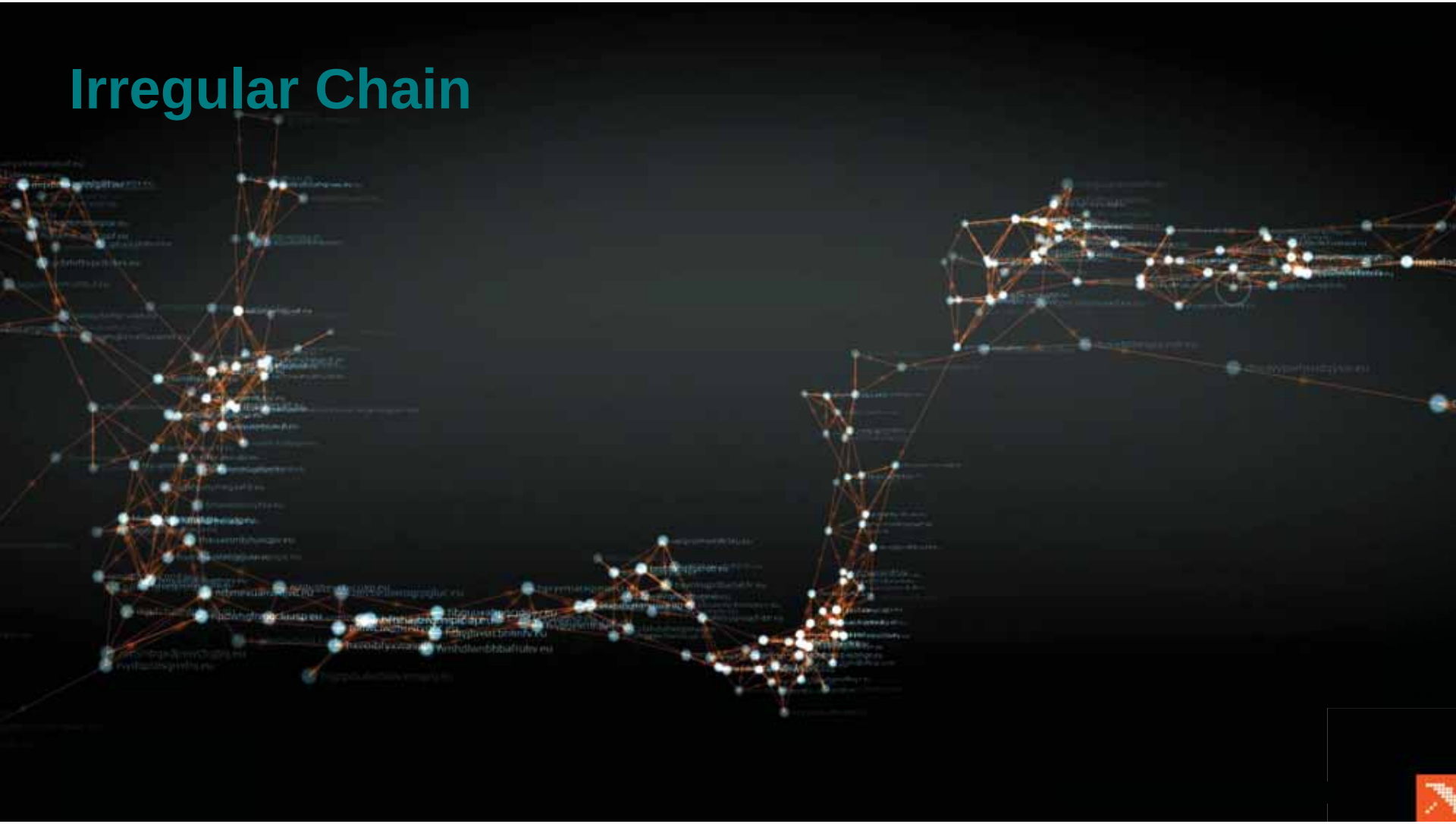
DGA-based Malware C2 Domains

- Ramnit
  - lyxbotuappfreadkfk.com
- Tinba
  - eersrjfoffvo.com, rlhupxlxoghh.com, vhtsiililikr.com
- Emotet
  - ywnjdkgrvivotium.eu
- Expiro
  - bpu1ilh.mazxceo.info

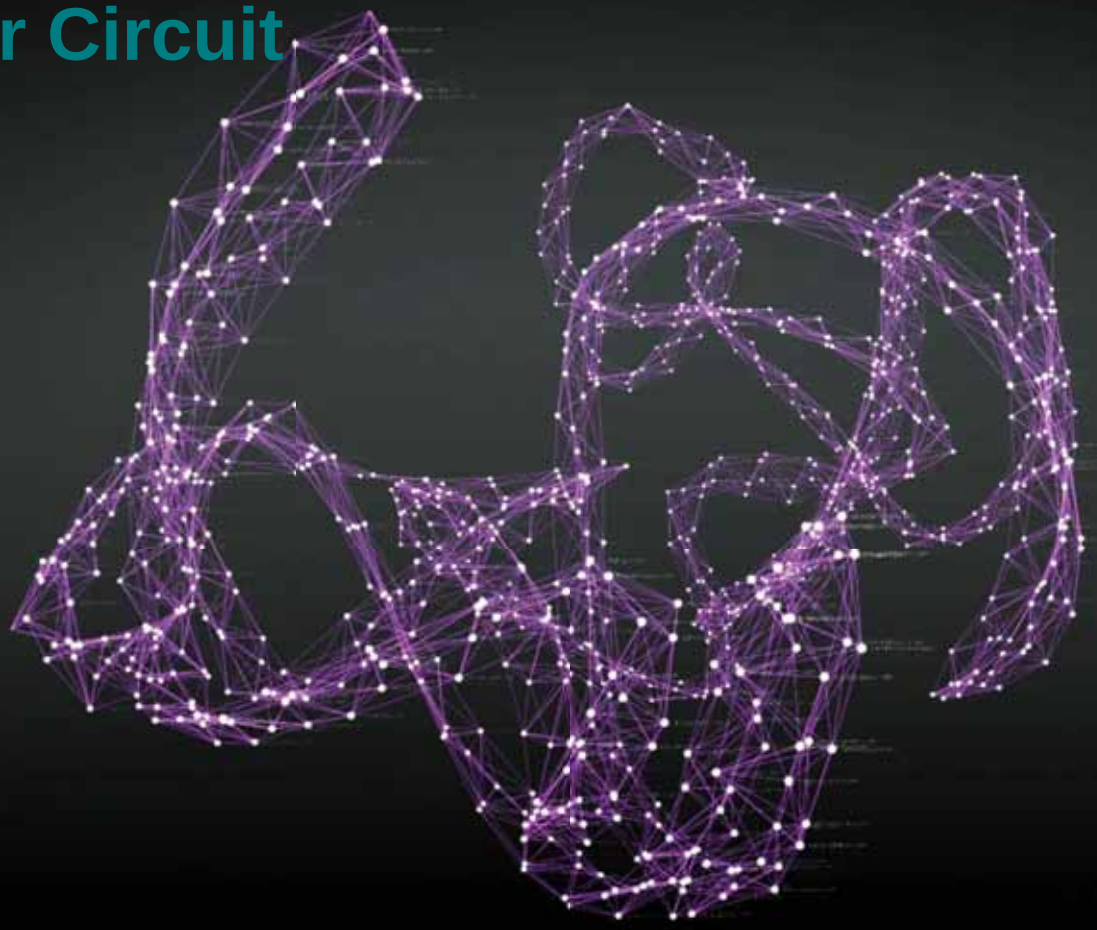
Regular Chain



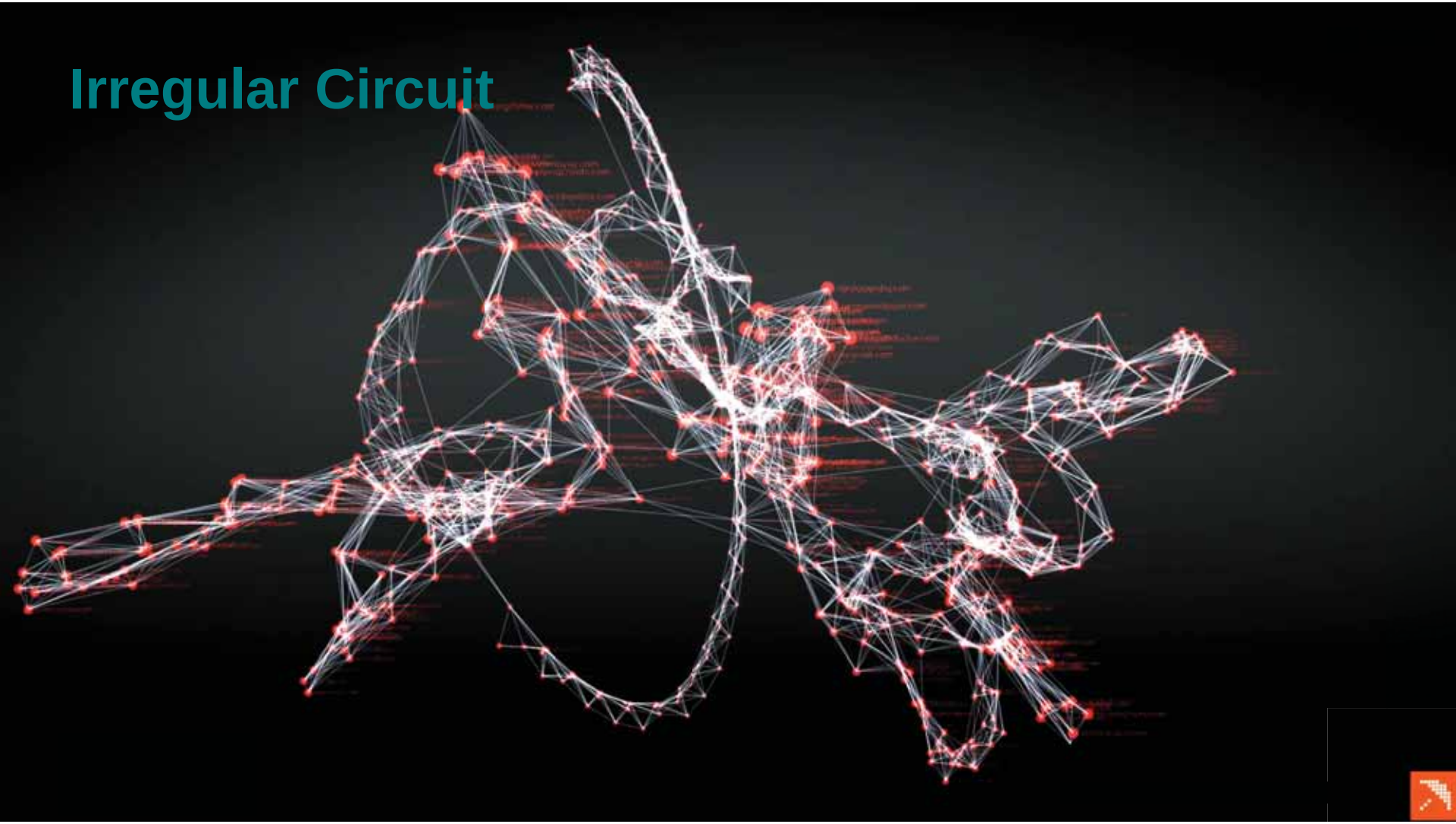
Irregular Chain



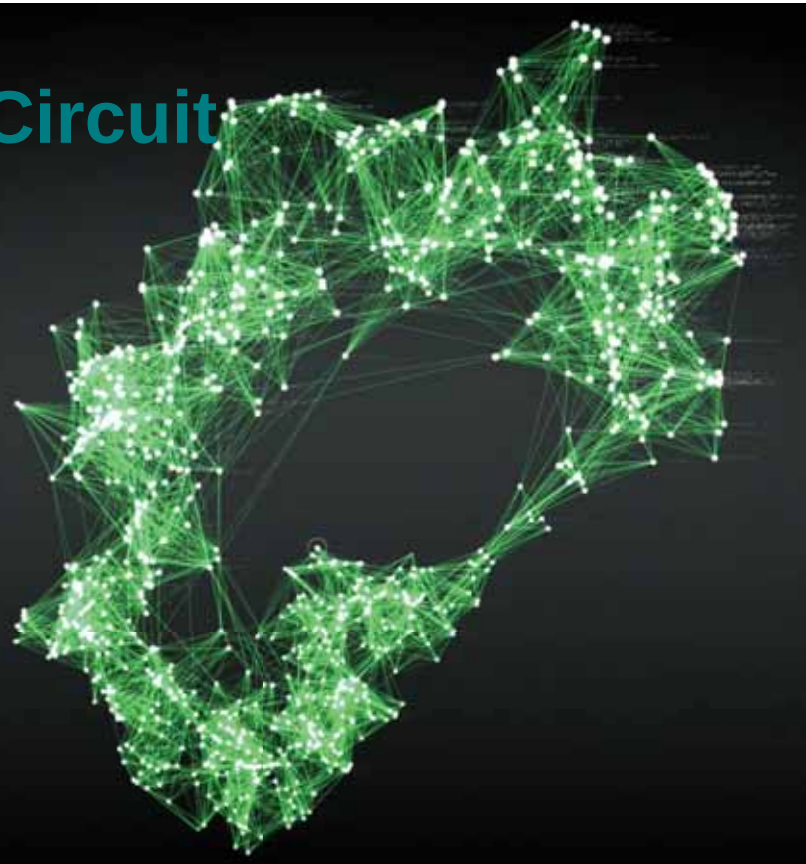
Regular Circuit



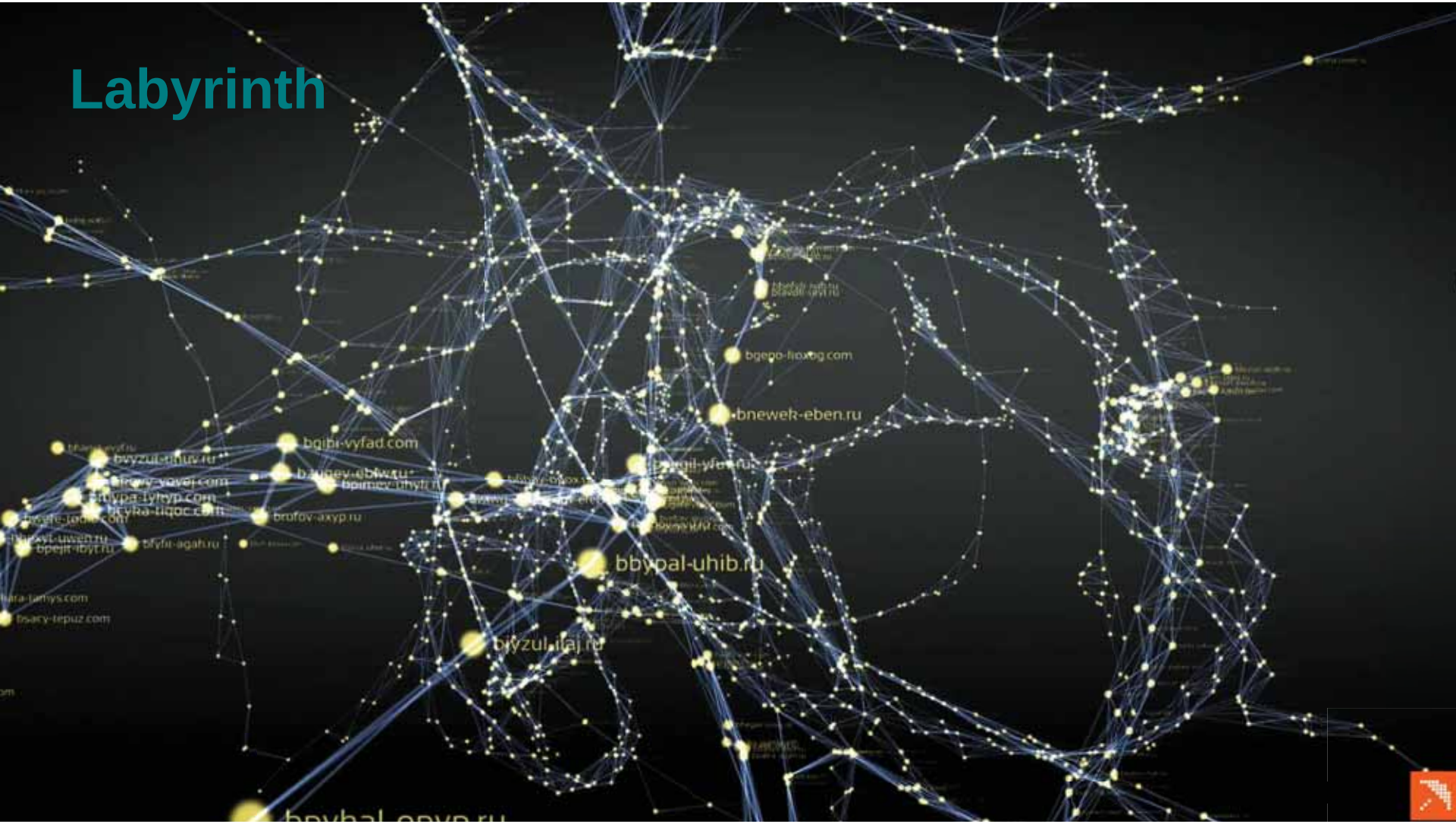
Irregular Circuit



Thick Irregular Circuit

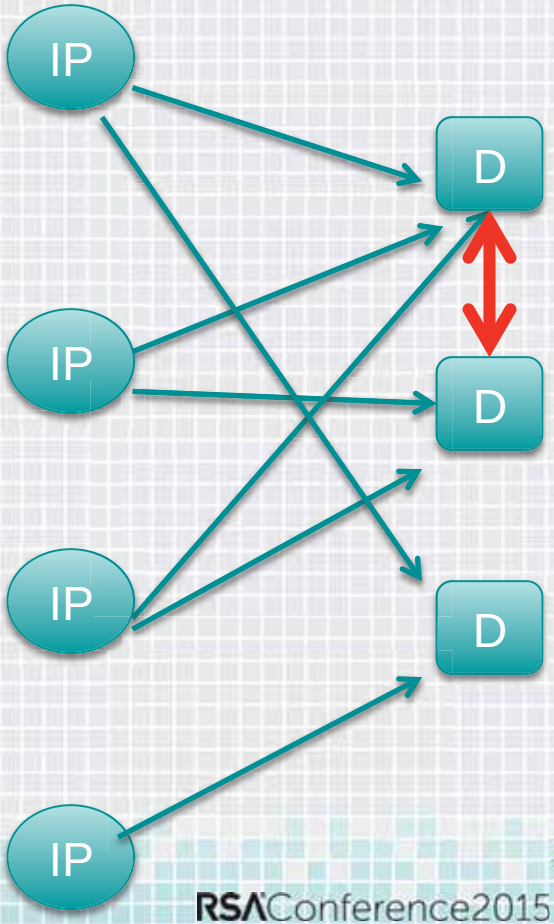


Labyrinth



Different Topologies?

- Consider co-occurrence time window
- Some clients are noisier than others
- Diversity in domain lookups of clients
- Nature of the DGA algorithm
  - Frequency,
  - Redundancy
  - Etc.



What Else Can I Use OpenGraphiti For?

Use OpenGraphiti…

- Provided data generation scripts
- File system
  - semanticnet/examples/fs_graph.py
- SHODAN query
  - semanticnet/examples/shodan_graph.py
- BRO IDS logs
  - semanticnet/examples/bro_graph.py

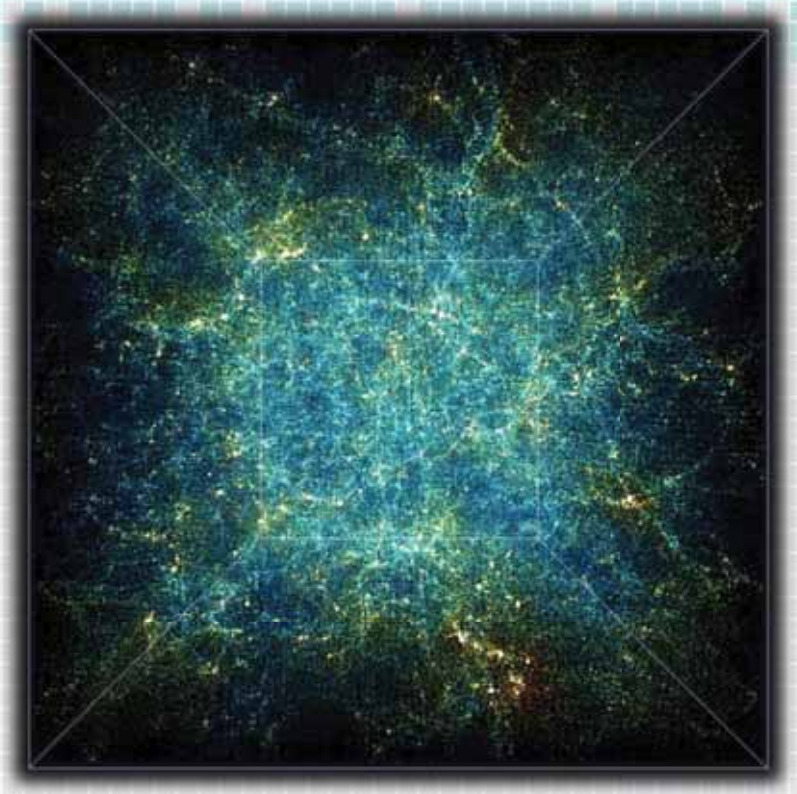
Use OpenGraphiti…

- Network packet captures
- IDS alerts
  - e.g. Snort, Bro, Suricata, etc.
- Environmental data
  - e.g. wind, water, earthquake, temperature, tide, soil statistics
- Odd data
  - e.g. Migratory patterns of the African and European coconut-laden swallow population

What's Next?

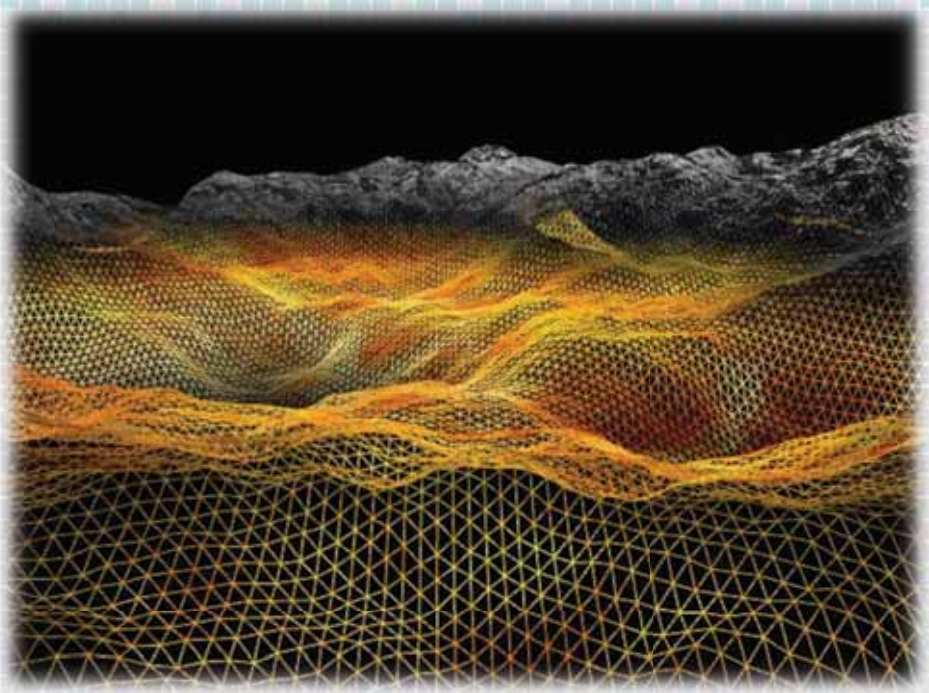
Smart Query Language

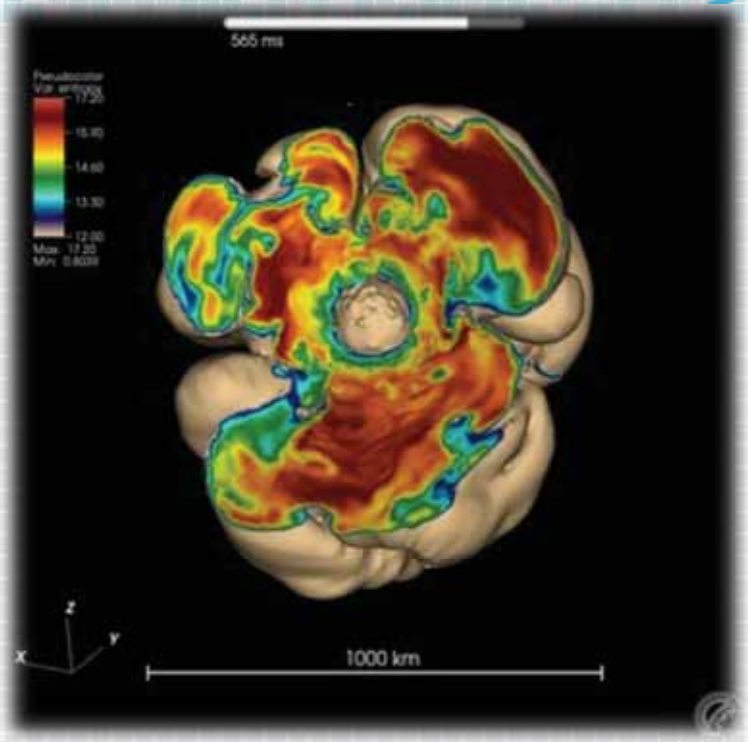
World View



Point Clouds

Time Series





Heightmaps

Volumes

Virtual Reality Experience

OpenGraphiti 1.0++

Apply Slide

- Data must be beautiful, interesting, and accurate to be useful
- Graphs are everywhere
  - Organize your data in graphs to quickly spot relationships and anomalies
- OpenGraphiti is
  - A free, Open Source, and awesome data visualization tool
  - Used to visualize any relational data as an interactive 2D or 3D model
  - http://www.opengraphiti.com/

RSAConference2015

San Francisco | April 20-24 | Moscone Center

SESSION ID: ECO-T08

Questions?

www.opengraphiti.com

Andrew Hay

Research Director
OpenDNS, Inc.
@andrewsmhay

Thibault Reuille

Sr. Security Researcher
OpenDNS, Inc.
@thibaultreuille