



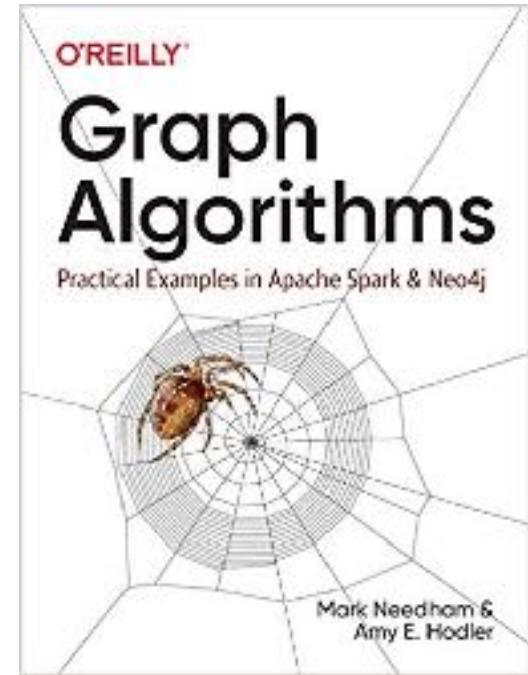
# Twitter Trolls Panama Papers

Panama Papers  
Paradise Papers

Justin Fine  
Westlake Village, CA  
June 2019

# Neo4j Graph Algorithms

The screenshot shows a blog post titled "Graph Algorithms in Neo4j: Graph Technology & AI Applications". The post discusses how graph technologies are being utilized for building intelligent applications, enabling more accurate predictions and faster decisions. It highlights various use cases such as recommendation engines, fraud detection, and machine learning. The post includes a call-to-action button "Download Now".



<https://neo4j.com/blog/graph-algorithms-neo4j-graph-technology-ai-applications/>

# Machine Learning and Graph Algorithms in Neo4j



## Pathfinding & Search

- Parallel Breadth First Search & DFS
- Shortest Path
- Single-Source Shortest Path
- All Pairs Shortest Path
- Minimum Spanning Tree
- A\* Shortest Path
- Yen's K Shortest Path
- K-Spanning Tree (MST)



## Centrality / Importance

- Degree Centrality
- Closeness Centrality
- CC Variations: Harmonic, Dangalchev, & Wasserman & Faust
- Betweenness Centrality
- Approximate Betweenness Centrality
- PageRank
- Personalized PageRank
- ArticleRank



## Community Detection

- Triangle Count
- Clustering Coefficients
- Connected Components (Union Find)
- Strongly Connected Components
- Label Propagation
- Louvain Modularity – 1 Step & Multi-Step
- Balanced Triad (identification)



## Similarity & ML Workflow

- Euclidean Distance
- Cosine Similarity
- Jaccard Similarity
- Overlap Similarity
- Random Walk
- One Hot Encoding

<https://neo4j.com/docs/graph-algorithms/3.5/>

Reference Implementations  
for Graph Embeddings  
(Node to Vector)  
• DeepGL  
• DeepWalk



# Justin Fine

Field Engineer , Neo4j

[Justin.Fine@neo4j.com](mailto:Justin.Fine@neo4j.com)

<https://www.linkedin.com/in/justin-fine-a7604523>

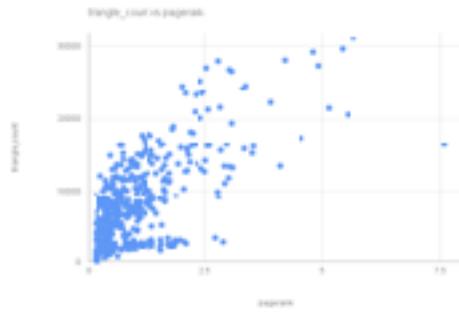
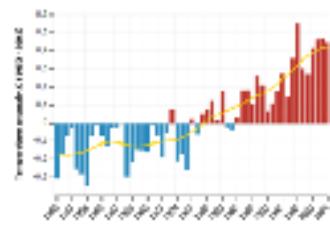
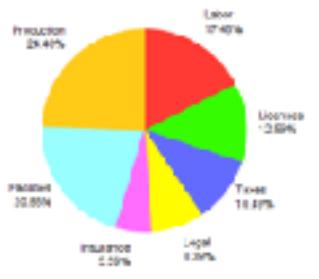
<https://github.com/JustinFineNeo/IntroToNeo4J>



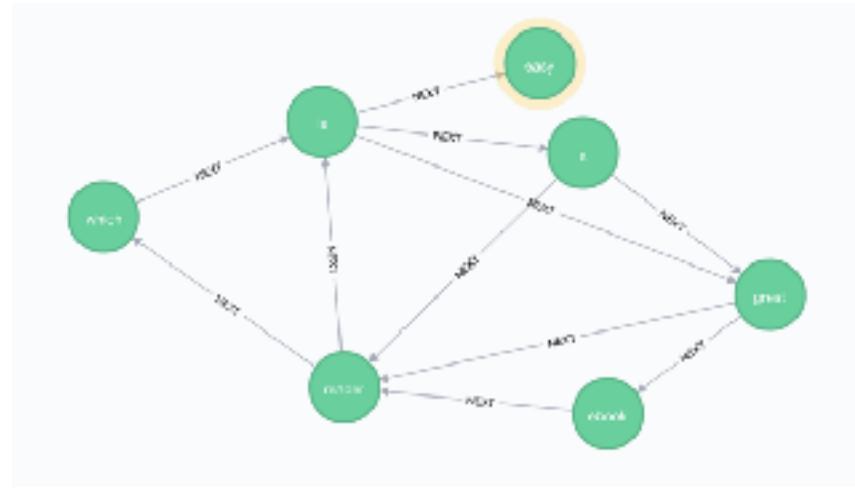
# What's a Graph Database???



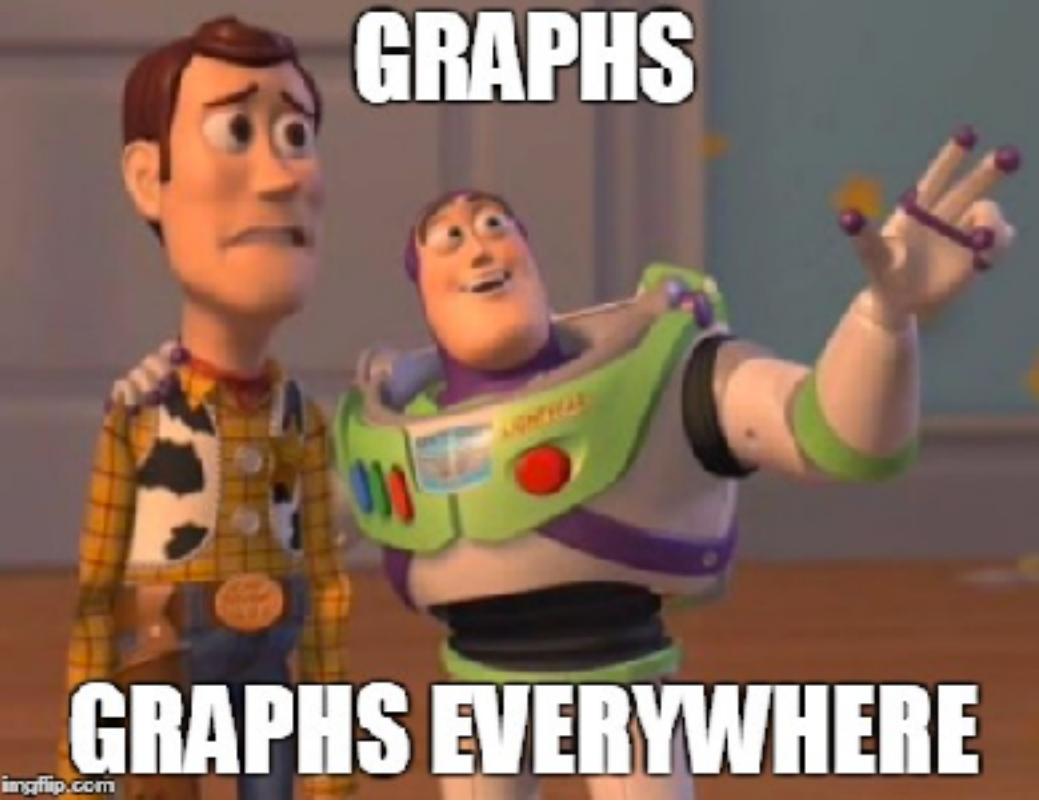
# Chart



# Graph

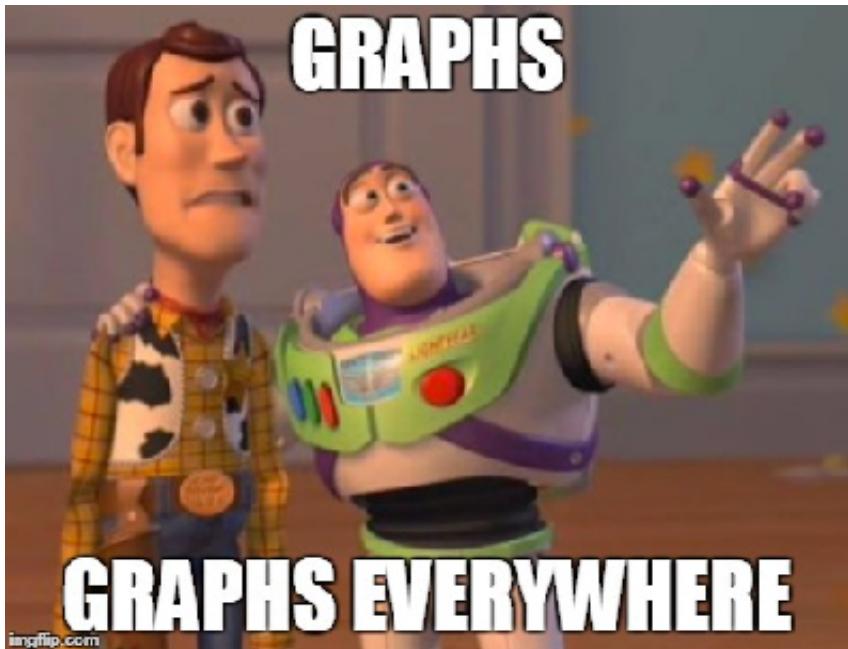


# GRAPHS

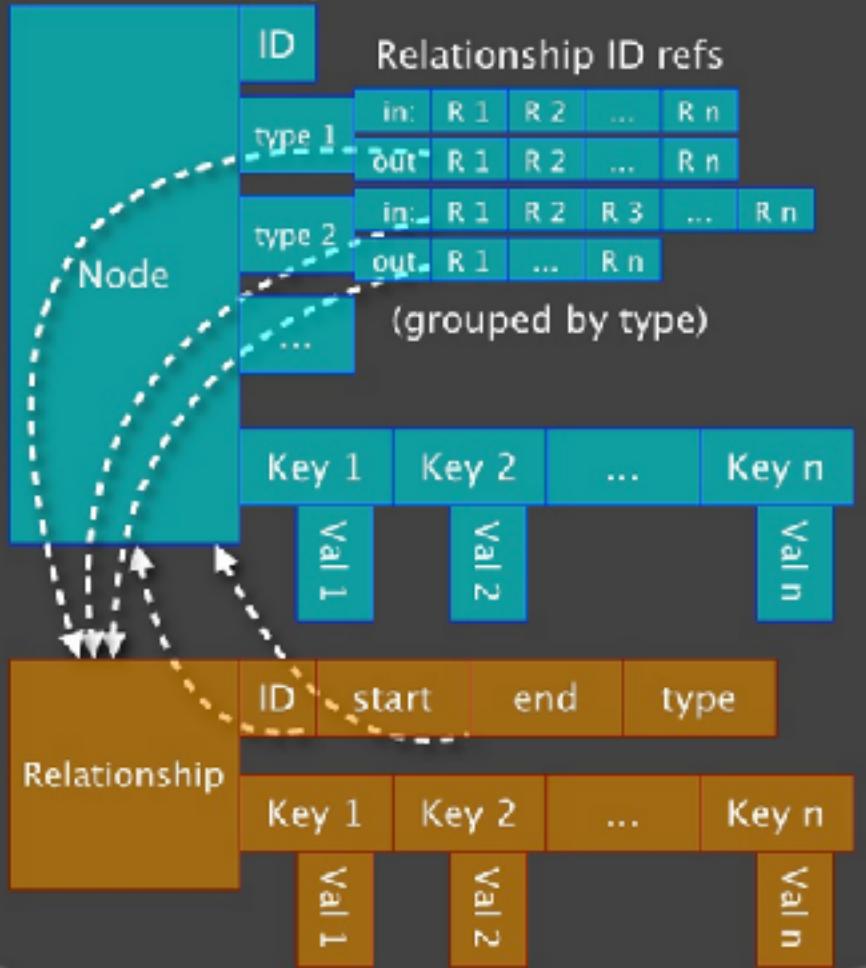


# Common Use Case

- Real time recommendations
- Fraud Detection
- Network & IT Management
- Social Networks
- Bill of Materials
- Knowledge Graphs
- Master Data Management
- Access Management
- Microservices Analysis
- IoT
- ...



## What we put in cache



## Neo4j Secret Sauce

- 1** Pointers instead of Lookups
- 2** Fixed Sized Records
- 3** “Joins” on Creation
- 4** Spin Spin Spin through this data structure

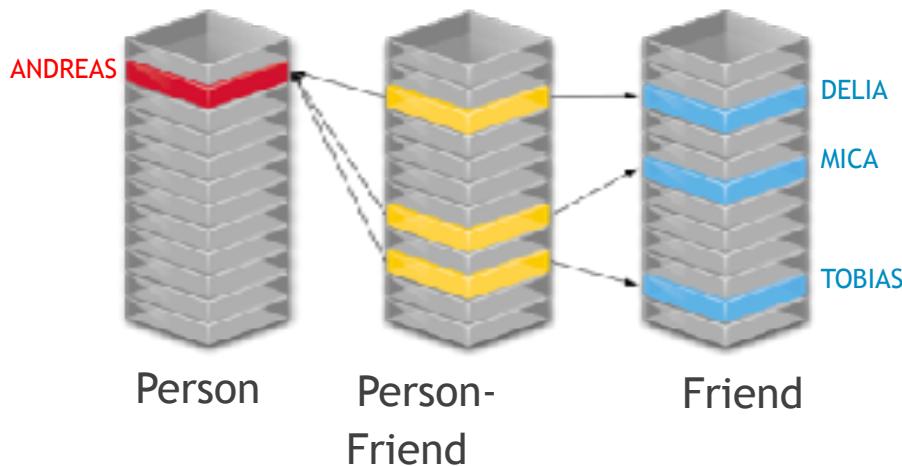
# **Neo4j Mission Statement**

**To help the world make sense of data**

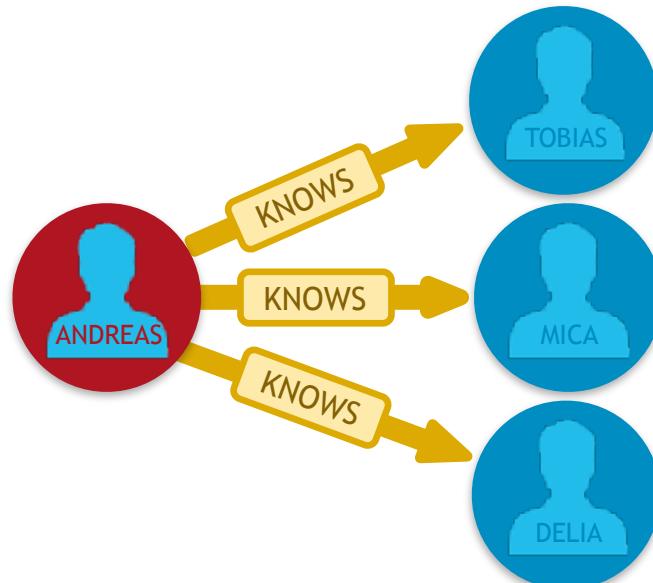
1 last\_name,first\_name,birthday,gender,type,state,district,party,url,address,phone,contact\_form,rss\_url,twitter,facebook,facebook\_id,youtube,youtube\_id,l  
2 Brown,Sherrod,1952-11-09,M,sen,OH,,Democrat,<http://www.brown.senate.gov>,713 Hart Senate Office Building Washington DC 20510,202-224-2315,<http://www.bn>  
3 Cantwell,Maria,1958-10-13,F,sen,WA,,Democrat,<http://www.cantwell.senate.gov>,511 Hart Senate Office Building Washington DC 20510,202-224-3441,<http://www.ca>  
4 Cardin,Benjamin,1943-10-05,M,sen,MD,,Democrat,<http://www.cardin.senate.gov>,589 Hart Senate Office Building Washington DC 20510,202-224-4524,<http://www.ca>  
5 Carper,Thomas,1947-01-23,M,sen,DE,,Democrat,<http://www.carper.senate.gov>,513 Hart Senate Office Building Washington DC 20510,202-224-2441,<http://www.ca>  
6 Casey,Robert,1960-04-13,M,sen,PA,,Democrat,<http://www.casey.senate.gov>,393 Russell Senate Office Building Washington DC 20510,202-224-6324,<http://www.ca>  
7 Corker,Bob,1952-06-24,M,sen,TN,,Republican,<http://www.corker.senate.gov>,425 Dirksen Senate Office Building Washington DC 20510,202-224-3344,<http://www.ca>  
8 Feinstein,Dianne,1933-06-22,F,sen,CA,,Democrat,<http://www.feinstein.senate.gov>,331 Hart Senate Office Building Washington DC 20510,202-224-3841,<https://>  
9 Hatch, Orrin,1934-08-22,M,sen,UT,,Republican,<https://www.hatch.senate.gov>,104 Hart Senate Office Building Washington DC 20510,202-224-5251,<http://www.ha>  
10 Klobuchar,Amy,1960-05-25,F,sen,MN,,Democrat,<http://www.klobuchar.senate.gov>,302 Hart Senate Office Building Washington DC 20510,202-224-3244,<http://www.ca>  
11 McCaskill,Claire,1953-07-24,F,sen,MU,,Democrat,<http://www.mccaskill.senate.gov>,738 Hart Senate Office Building Washington DC 20510,202-224-6154,<http://w>  
12 Menéndez,Robert,1954-01-01,M,sen,NJ,,Democrat,<http://www.menendez.senate.gov>,528 Hart Senate Office Building Washington DC 20510,202-224-4744,<http://w>  
13 Nelson,Bill,1942-09-29,M,sen,FL,,Democrat,<http://www.billnelson.senate.gov>,716 Hart Senate Office Building Washington DC 20510,202-224-5274,<http://www.ca>  
14 Sanders,Bernard,1941-09-08,M,sen,VT,,Independent,<http://www.sanders.senate.gov>,332 Dirksen Senate Office Building Washington DC 20510,202-224-5141,<http://w>  
15 Stabenow,Debbie,1956-04-29,F,sen,MI,,Democrat,<http://www.stabenow.senate.gov>,731 Hart Senate Office Building Washington DC 20510,202-224-4822,<http://w>  
16 Tester,Jon,1950-06-21,M,sen,MT,,Democrat,<http://www.tester.senate.gov>,311 Hart Senate Office Building Washington DC 20510,202-224-2644,<http://www.test>  
17 Whitehouse,Sheldon,1955-10-20,M,sen,RI,,Democrat,<http://www.whitehouse.senate.gov>,538 Hart Senate Office Building Washington DC 20510,202-224-2921,<http://w>  
18 Barrasso,John,1952-07-21,M,sen,WY,,Republican,<http://www.barrasso.senate.gov>,307 Dirksen Senate Office Building Washington DC 20510,202-224-6441,<http://w>  
19 Wicker,Roger,1951-07-05,M,sen,MS,,Republican,<http://www.wicker.senate.gov>,555 Dirksen Senate Office Building Washington DC 20510,202-224-6253,<http://w>  
20 Alexander,Lamar,1946-07-03,M,sen,TN,,Republican,<http://www.alexander.senate.gov>,455 Dirksen Senate Office Building Washington DC 20510,202-224-4944,<ht>  
21 Cochran,Thad,1937-12-07,M,sen,MS,,Republican,<http://www.cochran.senate.gov>,113 Dirksen Senate Office Building Washington DC 20510,202-224-5054,<http://w>  
22 Collins,Susan,1952-12-07,F,sen,ME,,Republican,<http://www.collins.senate.gov>,413 Dirksen Senate Office Building Washington DC 20510,202-224-2523,<http://w>  
23 Cornyn,John,1952-02-02,M,sen,TX,,Republican,<http://www.cornyn.senate.gov>,517 Hart Senate Office Building Washington DC 20510,202-224-2934,<http://www.co>  
24 Durbin,Richard,1944-11-21,M,sen,IL,,Democrat,<http://www.durbin.senate.gov>,711 Hart Senate Office Building Washington DC 20510,202-224-2152,<http://www.du>  
25 Enzi,Michael,1944-02-01,M,sen,WY,,Republican,<https://www.enzi.senate.gov>,379A Russell Senate Office Building Washington DC 20510,202-224-3424,<http://ww>  
26 Graham,Lindsey,1955-07-09,M,sen,SC,,Republican,<http://www.lgraham.senate.gov>,299 Russell Senate Office Building Washington DC 20510,202-224-5972,<http://w>  
27 Inhofe,James,1934-11-17,M,sen,OK,,Republican,<http://www.inhofe.senate.gov>,286 Russell Senate Office Building Washington DC 20510,202-224-4721,<http://w>  
28 McConnell,Mitch,1942-02-28,M,sen,KY,,Republican,<http://www.mcconnell.senate.gov>,317 Russell Senate Office Building Washington DC 20510,202-224-2541,<ht>  
29 Merkley,Jeff,1950-10-24,M,sen,OR,,Democrat,<http://www.merkley.senate.gov>,313 Hart Senate Office Building Washington DC 20510,202-224-3753,<http://www.m>  
30 Reed,John,1949-11-12,M,sen,RI,,Democrat,<http://www.reed.senate.gov>,728 Hart Senate Office Building Washington DC 20510,202-224-4842,<http://www.reed.sen>  
31 Risch,James,1943-05-03,M,sen,ID,,Republican,<http://www.risch.senate.gov>,403 Russell Senate Office Building Washington DC 20510,202-224-2752,<http://w>  
32 Roberts,Pat,1936-04-28,M,sen,KS,,Republican,<http://www.roberts.senate.gov>,189 Hart Senate Office Building Washington DC 20510,202-224-4774,<http://w>  
33 Sessions,Jefferson,1948-12-24,M,sen,AL,,Republican,<http://www.sessions.senate.gov>,326 Russell Senate Office Building Washington DC 20510,202-224-4124,I  
34 Shaheen,Jeanne,1947-01-20,F,sen,NH,,Democrat,<http://www.shaheen.senate.gov>,505 Hart Senate Office Building Washington DC 20510,202-224-2841,<http://w>  
35 Udall,Tom,1948-05-18,M,sen,NM,,Democrat,<http://www.tomudall.senate.gov>,531 Hart Senate Office Building Washington DC 20510,202-224-6821,<http://www.tu>  
36 Warner,Merk,1954-12-15,M,sen,VA,,Democrat,<http://www.warner.senate.gov>,475 Russell Senate Office Building Washington DC 20510,202-224-2023,<http://w>  
37 Gillibrand,Kirsten,1960-12-09,F,sen,NY,,Democrat,<http://www.gillibrand.senate.gov>,478 Russell Senate Office Building Washington DC 20510,202-224-4451,I  
38 Franken,Alan,1951-06-21,M,sen,MN,,Democrat,<http://www.franken.senate.gov>,389 Hart Senate Office Building Washington DC 20510,202-224-5541,<http://w>  
39 Coons,Chris,1953-09-09,M,sen,DE,,Democrat,<http://www.coons.senate.gov>,1274 Russell Senate Office Building Washington DC 20510,202-224-5042,<http://w>  
40 Manchin,Joe,1947-08-24,M,sen,WV,,Democrat,<http://www.manchin.senate.gov>,386 Hart Senate Office Building Washington DC 20510,202-224-3954,<http://w>  
41 Aderholt,Robert,1955-07-22,M,rep,AL,4,Republican,<https://aderholt.house.gov>,235 Cannon HOB; Washington DC 20515-0104,202-225-4876,<http://aderholt.house.gov>  
42 Amash,Justin,1960-04-10,M,rep,MI,3,Republican,<http://amash.house.gov>,114 Cannon HOB; Washington DC 20515-2203,202-225-3001,<https://amash.house.gov/con>

# Relational Versus Graph Models

Relational Model



Graph Model



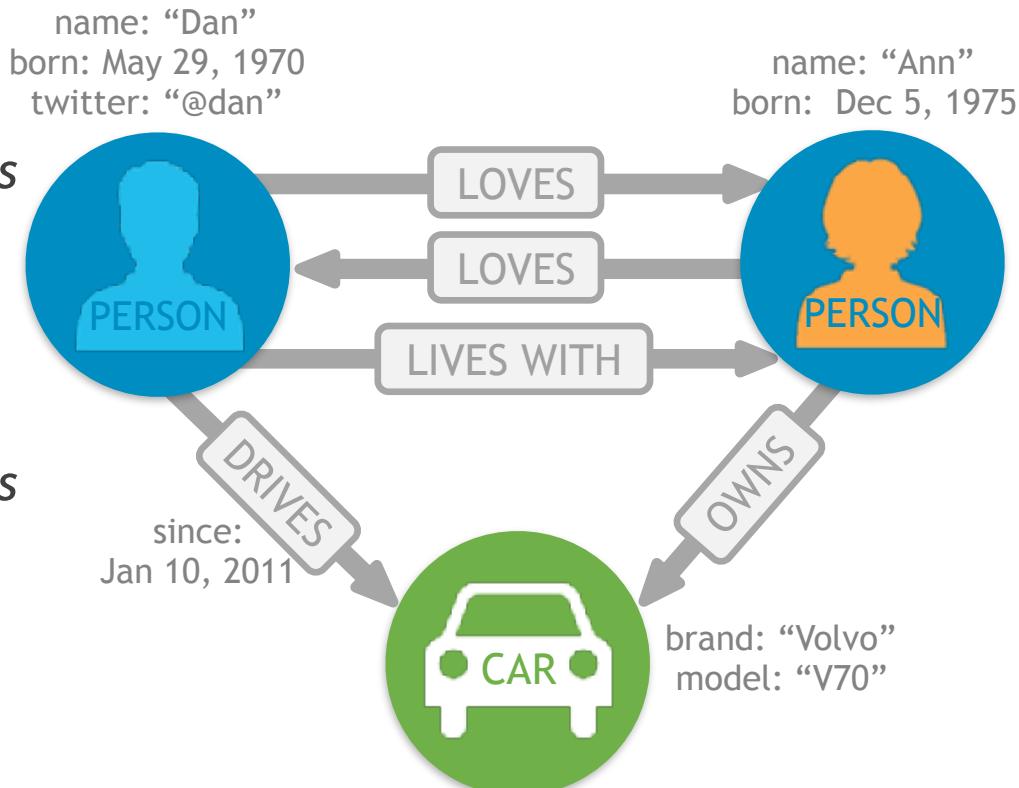
# Property Graph Model Components

## Nodes

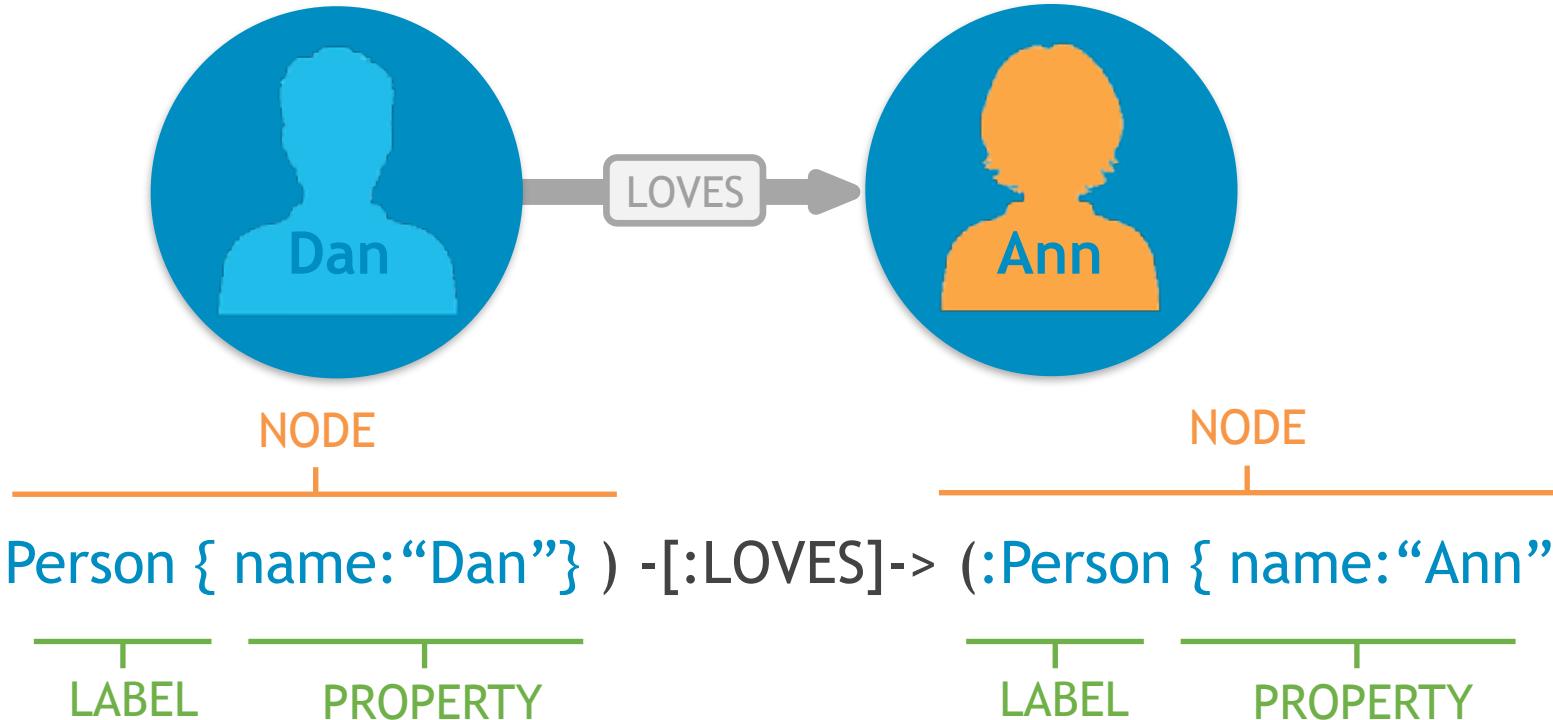
- The objects in the graph
- Can have name-value *properties*
- Can be *labeled*

## Relationships

- Relate nodes by type and direction
- Can have name-value *properties*



# Cypher: Powerful and Expressive Query Language



# Cypher Query Language

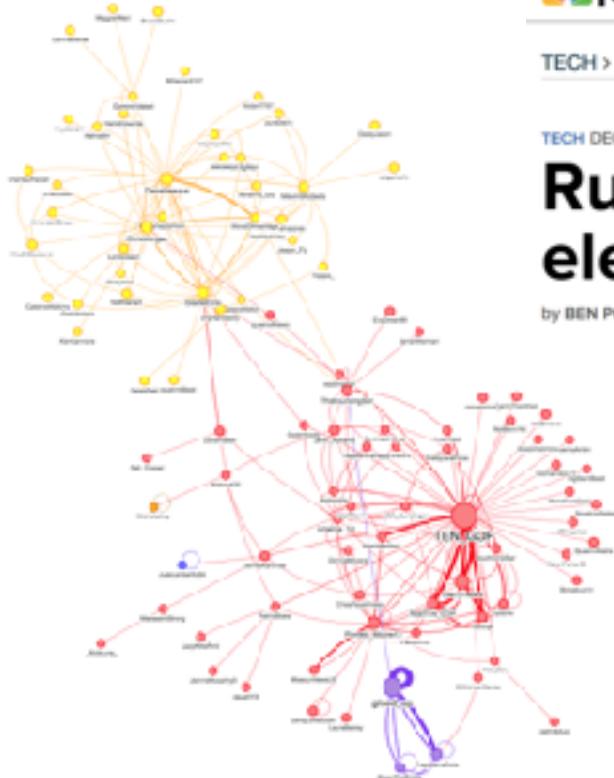
SQL for graphs

# Twitter, Russia, Trolls, Oh my!

TECH DEC 20 2017, 11:11 AM ET

# Russian trolls went on attack during key election moments

by BEN POPKEN



```
1 MATCH (tr:Troll)-[:POSTED]->(tw:Tweet) WITH tr, tw
2 OPTIONAL MATCH (tw)-[:RETWEETED]-(rt:Tweet)
3 OPTIONAL MATCH (tw)-[:IN_REPLY_TO]-(irp:Tweet)
4 RETURN distinct tr.screen_name as screen_name, count(tw) as totalTweets,
5 count(rt) as totalRetweets, count(irp) as totalReplies,
6 (count(tw) - (count(rt) + count(irp))) as originalContent
7 ORDER BY totalTweets DESC;
```

# Russia Twitter Trolls



- **2752 Twitter accounts tied to Russia's Internet Research Agency**
- Accounts suspended by Twitter
  - Data deleted
- **What were they tweeting about?**

user id	handle
905874659358453780	10_gop
2528776085	1488reasons
2535564756	1D_Nicole_
79793901	1DRussianFNDM
1860330774	1Erik_Lee
638742761516891041	1orenatawa1
2537164155	2auregbacon
623484279767294720	2ndHalfOnion
80712878	2olvbaker
80723578	30ToMarsFandom
3644473036	489Jisalge
2281001126	4clainsavara
748870306260647168	4ever1937
4038537452	4MySquad
713885792550545760	606stavarogers
2534876524	6Cruz
3038735149	71bilajamitII
2533199049	706shawn
2919792207	__Judith__0_
2556910923	50phlia
28854866301	_alshaw
2544963209	Amy_McAann_
2763362910	AnnaSwanson
2540386370	_Decker_Gerald
2428927403	_beglow
2559290335	_Ben_Santos

# Wayback API



BBR ID	handle
905874650356453760	10_gop
2538776085	146seasons
2532554795	1D_Nicole
797039901	1DReynoldsMEMU
1993230774	1Erik_Lee
839742781515981041	1Ivanashva1
2937191150	2auragleson
623434279787294720	2ndLeftOnion
80712873	2efverbeker
60720570	307MiversRandom
3544475035	459Justiga
2281091126	Ackarocvara
748870305380648168	Aever1907
403537482	4WYSpacel
7138887055600460760	6890devonge05
2534875985	6Dnuz
3936735149	71bilaljamil
2933189949	760shawm
2915792207	__Jacin__D_
2556810923	__Sophia__
2593460901	_alishaw
2544663209	Amy_Mocean
2753582910	AnnaSvanson
25403866370	Bekley_Gerali
2426327403	bingov
2592900395	Ben_Santos
3432464074	Billy_Moyer_
2583885248	Edward_Connor_
2914885389	Gazelle_Popo_
2592014210	GeorgeGschutz_
2445237434	GregorVan
2922070917	Howard_Good_



[http://archive.org/wayback/available?url=http://twitter.com/TEN\\_GOP](http://archive.org/wayback/available?url=http://twitter.com/TEN_GOP)



```
{  
    "url": "https://twitter.com/TEN_GOP",  
    "archived_snapshots": [  
        {  
            "closest": {  
                "status": "200",  
                "available": true,  
                "url": "https://web.archive.org/web/20170818065925/https://twitter.com/TEN_GOP",  
                "timestamp": "20170818065926"  
            }  
        }  
    ]  
}
```

# Scraping Internet Archive...

Tennessee Mengelwest Semua  
Tennessee @TEN\_GOP · 11  
About a week ago a horde of Moroccans landed on the beach in Spain. This week a terror attack by a Moroccan in #Barcelona.



0 13 0

```
{"suggestions_details":{}, "tweet_ids": "898272614224769624", "scribe_component": "tweet"}>  
::before  
+div class="tweet js-stream-tweet js-actionable-tweet js-profile-popup-actionable dismissible-content original-tweet js-original-tweet  
  
has-cards has-content  
+ data-tweet-id="898272614224769624" data-item-id="898272614224769624" data-permalink-path="/TBN_GOP/status/898272614224769624" data-conversation-id="898272614224769624" data-tweet-source="898272614224769624" data-retweeter="TEN_GOP" data-screener-name="TEN_GOP" data-name="Tennessee" data-user-id="42200229004" data-you-follow="false" data-follows-you="false" data-you-block="false" data-reply-to-user-json="{}" id_str="4224729944" screen_name="TBN_GOP" name="Tennessee" emojiified_name="Tennessee" emojified_text_as_html="Tennessee"))" data-disclosure-type="data-has-cards"="true">> = 50  
::before  
+div class="context"></div>  
+div class="content">  
+div class="stream-item-header"></div>  
+div class="js-tweet-text-container">  
+p class="TweetTextSize TweetTextSize--normal js-tweet-text tweet-text" lang="en" data-aria-label-part="0">  
"About a week ago a horde of Moroccans landed on the beach in Spain. This week a terror attack by a Moroccan in "  
+a href="https://web/26178865826/https://twitter.com/hashtag/Barcelona?src=hash" data-query-source="hashtag_click" class="twitter-hashtag pretty-link is-new" dir="ltr">></a>
```

# Scraping Internet Archive...

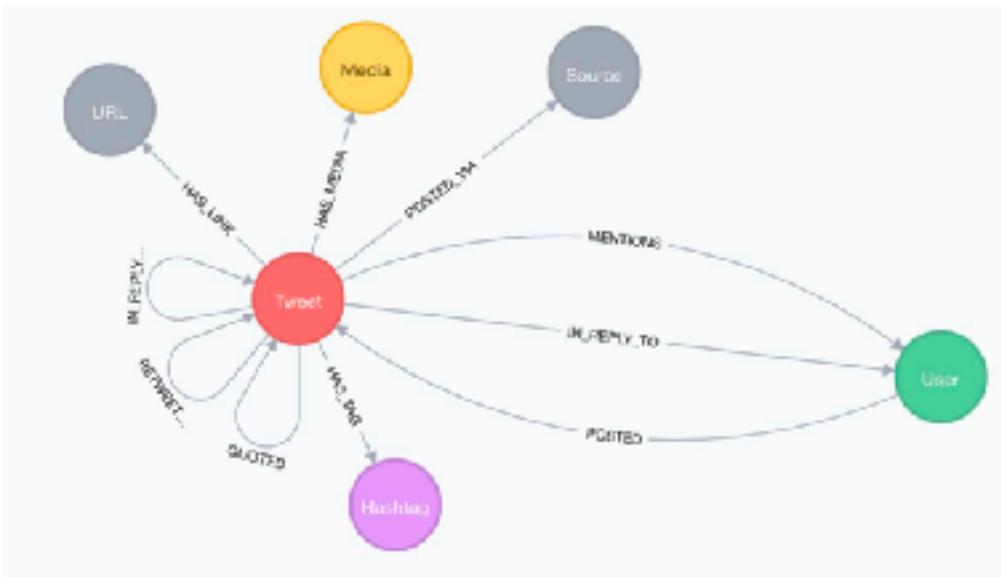


```
1 WITH $tweetArr AS tweets
2 UNWIND tweets AS tweet
3
4 MERGE (u:User {user_id: tweet.user_id})
5 ON CREATE SET u.screen_name = tweet.screen_name
6
7 MERGE (t:Tweet {tweet_id: tweet.tweet_id})
8 ON CREATE SET t.text = tweet(tweet_text)
9
10 MERGE (u)-[:POSTED]->(t)
11
12 FOREACH (ht IN tweet.hashtags |
13   MERGE (h:Hashtag {tag: ht.tag })
14   ON CREATE SET h.archived_url = ht.archived_url
15   MERGE (t)-[:HAS_TAG]->(h)
16 )
17
18 FOREACH (link IN tweet.links |
19   MERGE (l:Link {url: link.url})
20   ON CREATE SET l.archived_url = link.archived_url
21   MERGE (t)-[:HAS_LINK]->(l)
22 )
23
```

# Scraping Internet Archive...

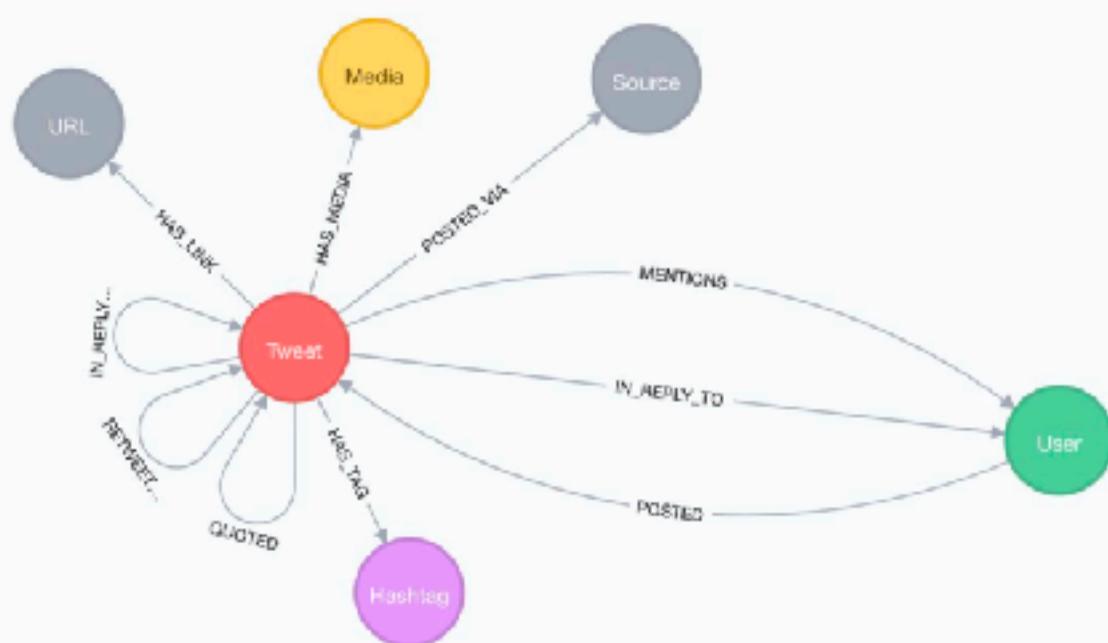


```
1 MERGE (User:User {id: 0}) AS User
2 MERGE (User:User {id: 45}) AS tweet
3
4 MERGE (t:User {user_id: tweet.user_id})
5 ON CREATE SET t.username = tweet.username
6
7 MERGE (t:Tweet {tweet_id: tweet.id})
8 ON CREATE SET t.text = tweet.text,
9     t.username = tweet.username
10
11 MERGE (t:REPO) ID t
12
13 FOREACH (l IN tweet.entities) |
14     MERGE (l:Hashtag {tag_id: l.tag_id})
15     ON CREATE SET l.archived_url = t.archived_url
16     MERGE (l:Hashtag {tag_id: l.tag_id})
17
18
19 FOREACH (link IN tweet.links) |
20     MERGE (l:Link {link_id: link.link_id})
21     ON CREATE SET l.archived_url = link.archived_url
22     MERGE (l:Hashtag {link_id: l.link_id})
23
24
```



# SUCCESS!!

345k Tweets, 41k Users (454 Russian Trolls)



Now what?

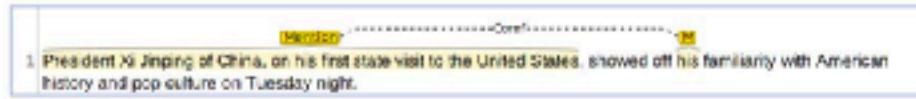
# Common NLP Tasks

- Language detection
- Part of speech tagging
- Word similarities
- Keyword extraction
- Topic extraction
- Concept hierarchy
- Sentiment analysis
- Entity extraction
- Entity merging
- Word associations
- Summarization
- Opinion mining

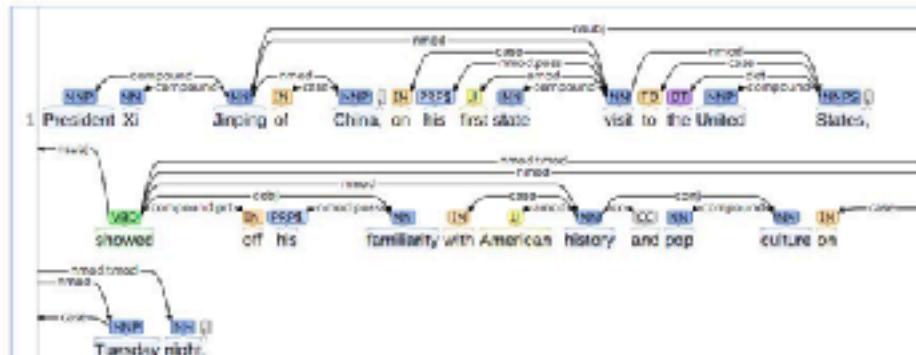
## Named Entity Recognition:



## Coreference:



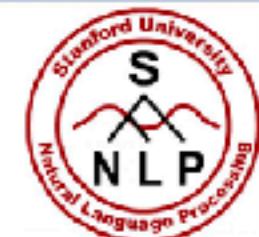
## Basic Dependencies:



NLTK  
Textblob  
Polyglot  
TextRank



ConceptNet

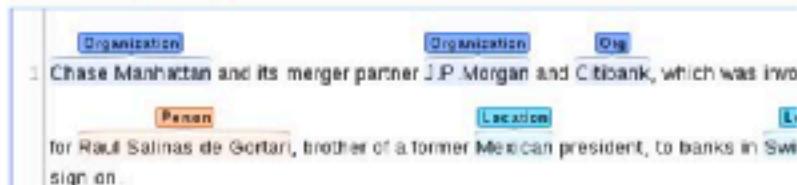


An open, multilingual knowledge graph

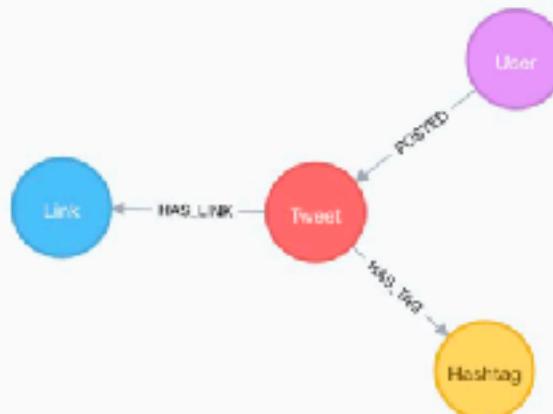
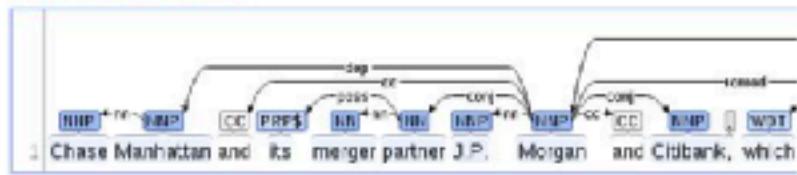
# NLP w/ Graph Databases

## Annotations

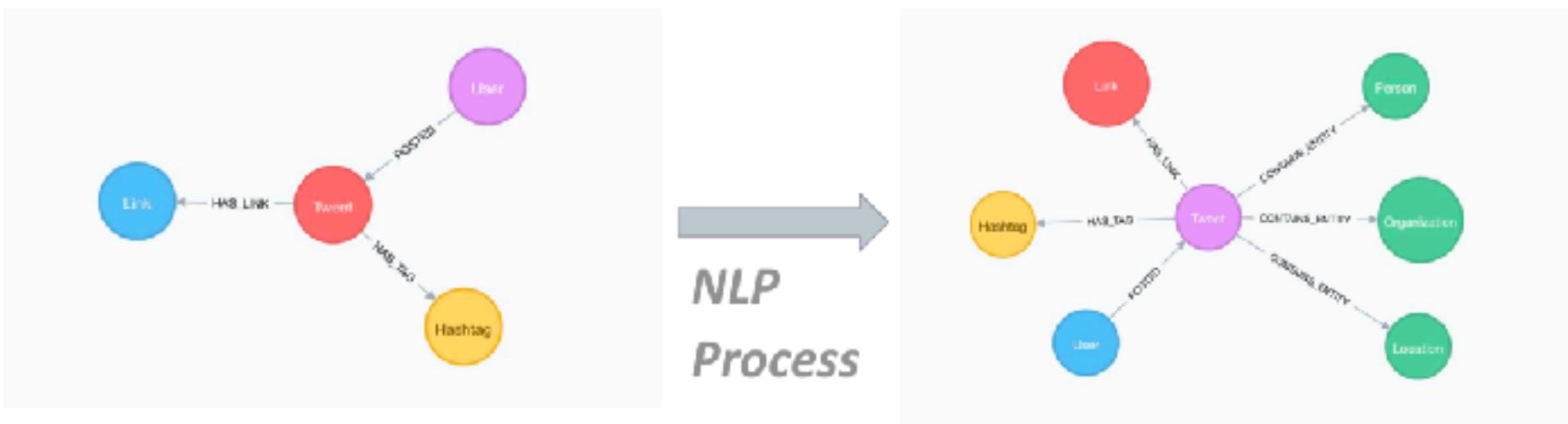
### Named Entity Recognition:



### Basic dependencies:



# NLP w/ Graph Databases *Annotations*



# How to combine open source NLP tools w/ Neo4j?



## GraphAware Natural Language Processing

build passing This [Neo4j](#) plugin offers Graph Based Natural Language Processing capabilities.

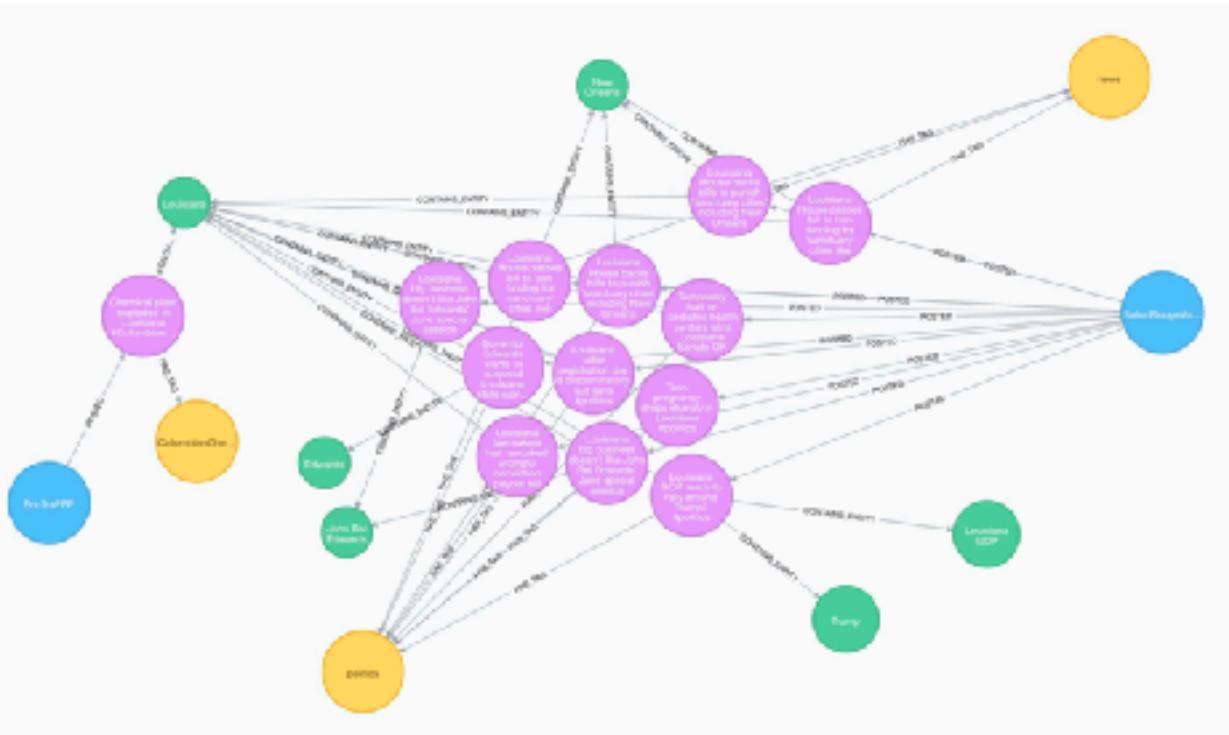


```
1 MATCH (tw:Tweet {lang: "en"})
2 CALL ga.nlp.annotate({text: tw.text, id: id(tw)})
3 YIELD result
4 MERGE (tw)-[:HAS_ANNOTATED_TEXT]->(result)
5 RETURN count(result)
```

<https://github.com/graphaware/neo4j-nlp>

PEOPLE MENTIONED		LOCATIONS MENTIONED		ORGANIZATIONS MENTIONED	
Person	Num	Location	Num	Organization	Num
"donald trump"	6675	"white house"	780	"white house"	760
"hillary clinton"	4702	"new york"	232	"new york"	232
"bill clinton"	900	"united states"	188	"north carolina"	138
"mike pence"	370	"milwaukee"	142	"hillary's health"	106
"tim kaine"	370	"north carolina"	138	"democratic party"	106
"michele obama"	358	"munich"	126	"supreme court"	100
"bernie sanders"	238	"saudi arabia"	104	"republican party"	98
"barack obama"	190	"baltimore"	104	"state department"	92
"rudy giuliani"	188	"philippines"	84	"democratic national convention"	84
"david darier"	170	"paris"	84	"gop convention"	76
"ted cruz"	164	"north korea"	78	"trump tower"	68
"ben carson"	163	"st louis"	76	"trump university"	58
"sheriff clarke"	160	"pittsburgh"	68	"justice department"	56
"newt gingrich"	155	"trump tower"	68	"islamic state"	52
"kopnews obama"	128	"utah"	68	"secret service"	46
"george zones"	114	"baton rouge"	56	"walk of fame"	44
"colin powell"	108	"middle east"	56	"house gop"	40
"leopoldo lópez"	88	"new york city"	56	"air force"	40
"trump foundation"	86	"laos"	52	"rust belt"	36
"paul ryan"	85	"atlantic city"	44	"china"	26
"chelsea clinton"	84	"los angeles"	44		
"leopoldo obama"	69				
"seeger"	78				

```
1 // Tweets, hashtags, and entities around Louisiana
2 MATCH (u:User)-[:POSTED]->(t:Tweet)-[:CONTAINS_ENTITY]->(l:Location {name: "Louisiana"})
3 OPTIONAL MATCH (t)-[:HAS_TAG]->(ht:Hashtag)
4 OPTIONAL MATCH (t)-[:CONTAINS_ENTITY]->(e)
5 RETURN *
```



# Your typical American Citizen?



@LeroyLovesUSA

Cleveland Online

@OnlineCleveland

Breaking news, weather, traffic  
and more for Cleveland. DM us  
anytime. RTs not endorsements

City of Cleveland, USA

@ClevelandOnline

# Your typical Local News Publication?



Tennessee

@TEN\_GOP

Unofficial Twitter of Tennessee  
Republicans. Covering breaking  
news, national politics, foreign  
policy and more. #MAGA #2A

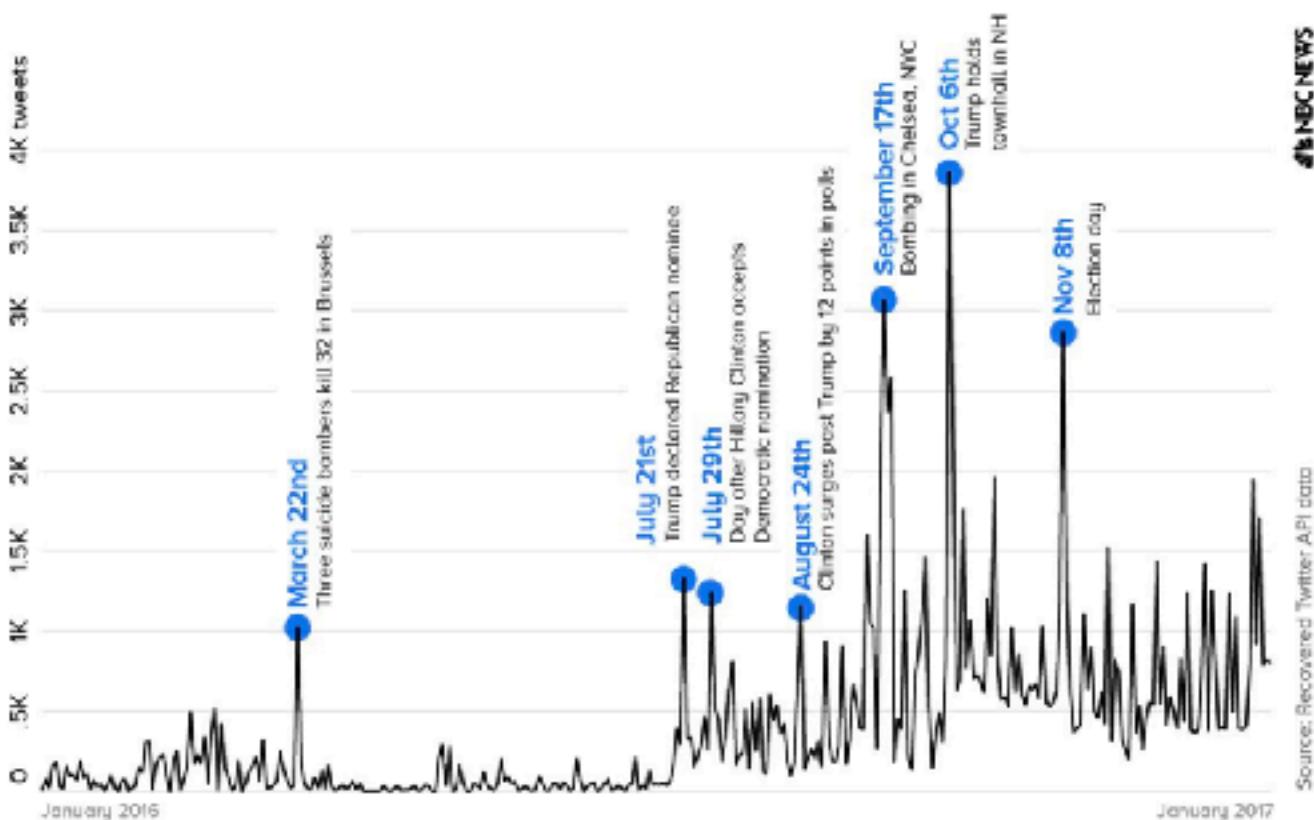
@TEN\_GOP

# Your typical Local Political Party?

```
1 MATCH
2 (u:User {screen_name: "LeroyLovesUSA"})-[:POSTED]->(t:Tweet)-[:HAS_TAG]->(ht:Hashtag {key: "thanksobama"})
3 RETURN *
```

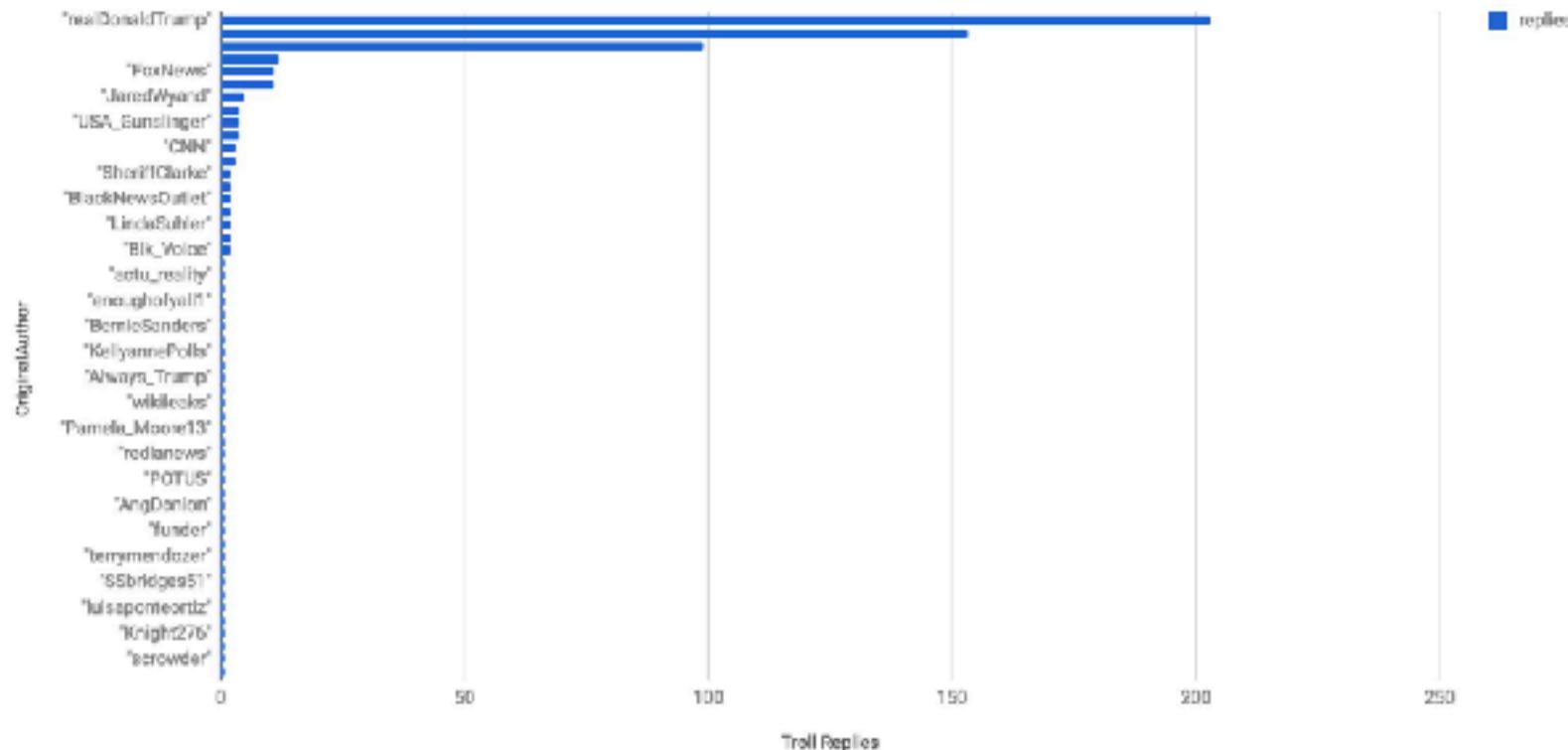


## Russian Troll Volume Spiked During 2016 Campaign Events

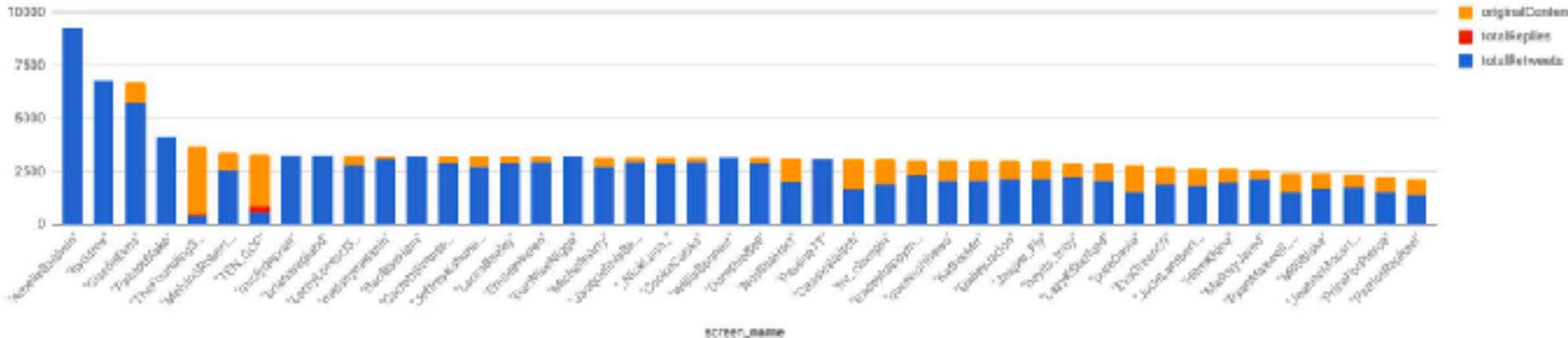


# NBC NEWS

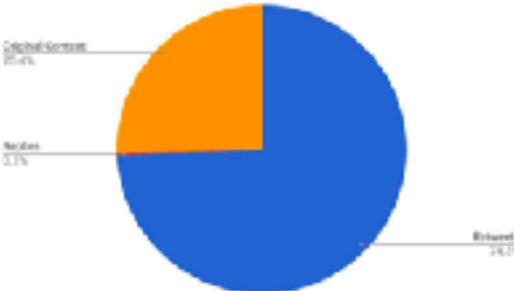
## Replies vs. Original Author



totalRetweets, total Replies and originalContent

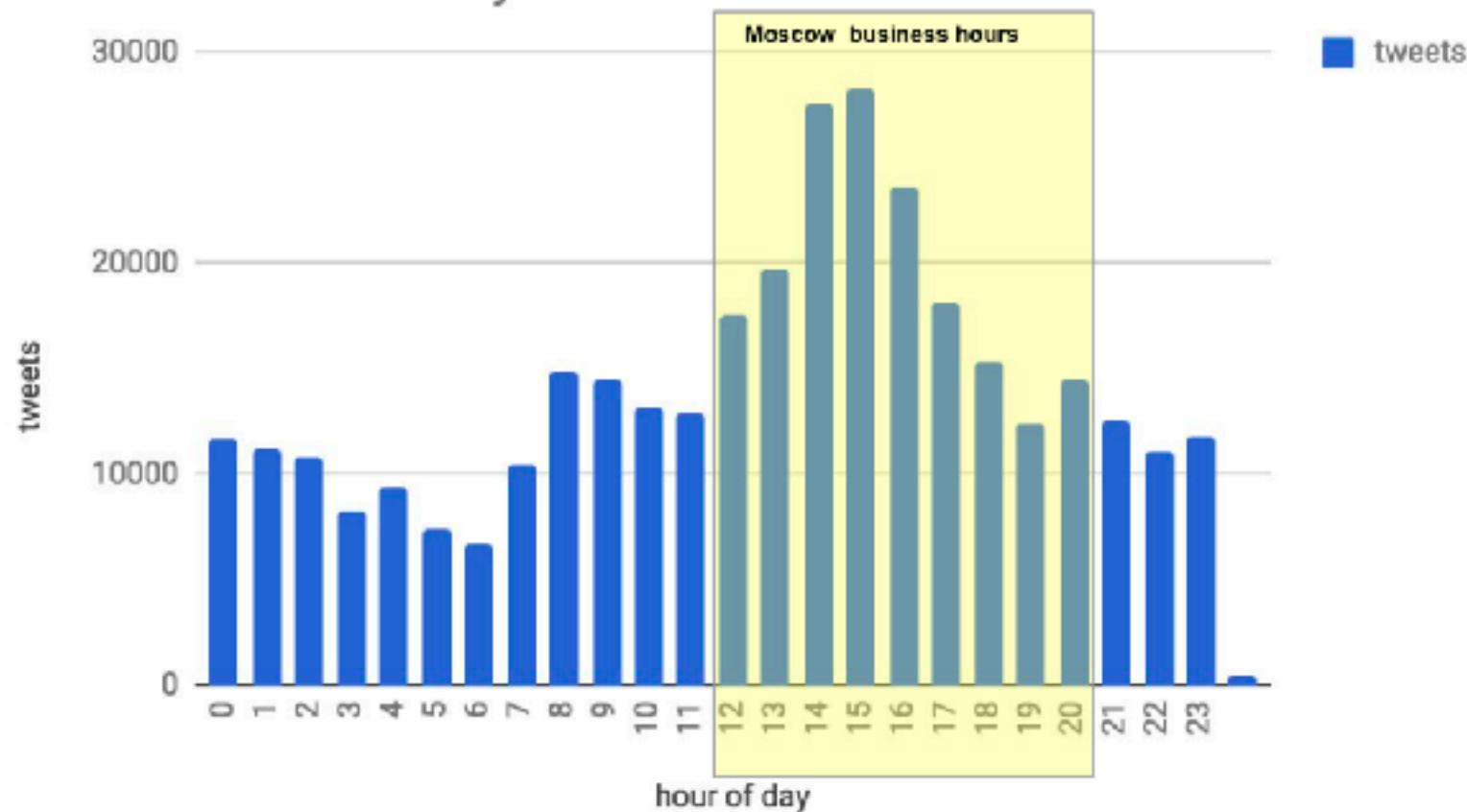


### Summary Troll Behavior



```
1 MATCH (tr:Troll)-[:POSTED]->(tw:Tweet) WITH tr, tw
2 OPTIONAL MATCH (tw)-[:RETWEETED]-(rt:Tweet)
3 OPTIONAL MATCH (tw)-[:IN_REPLY_TO]-(irp:Tweet)
4 RETURN distinct tr.screen_name as screen_name, count(tw) as totalTweets,
5      count(rt) as totalRetweets, count(irp) as totalReplies,
6      (count(tw) - (count(rt) + count(irp))) as originalContent
7 ORDER BY totalTweets DESC;
```

## tweets vs. hour of day



Your typical American Citizen?



@ZeroLoveUSA

Cleveland Online

@OnlineCleveland

Breaking news, weather, traffic  
and more for Cleveland. DM us  
anytime. RTs not endorsements

City of Cleveland, USA

@ClevelandOnline

Your typical Local Political Party?

# Russian

Your typical Local News Publication?



@TEN\_GOP

# **Panama Papers:**

**How Neo4j Bummed out some very high up  
people**

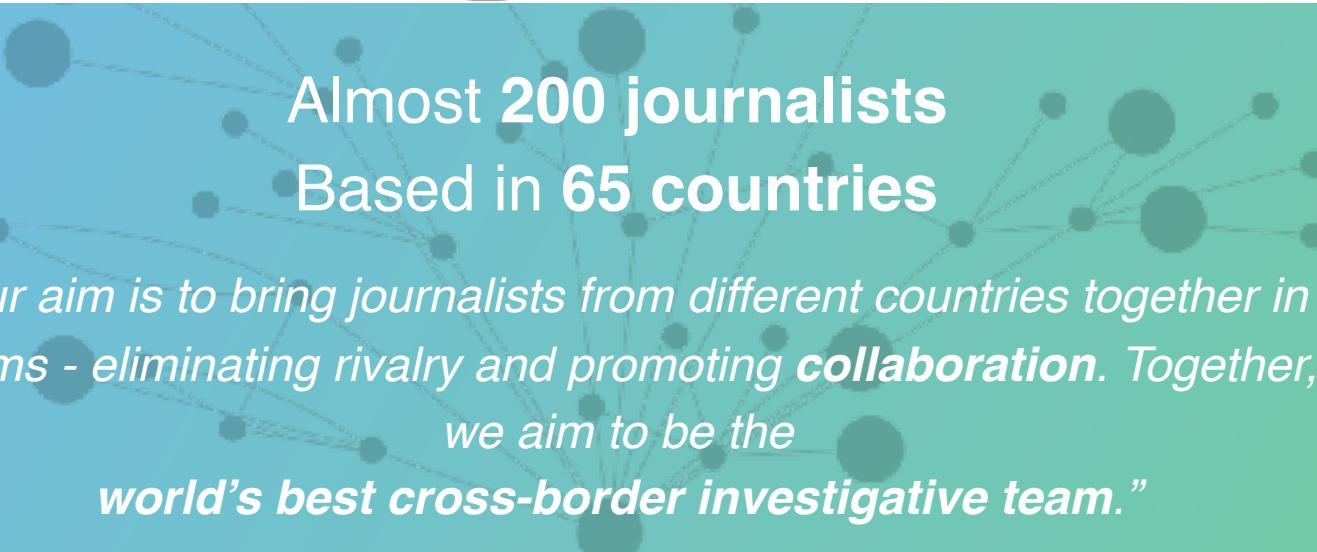
# THE PANAMA PAPERS

Politicians, Criminals and the Rogue Industry That Hides Their Cash





# The International Consortium of Investigative Journalists



A network diagram composed of numerous small, semi-transparent circles of varying sizes and shades of blue and grey, connected by thin grey lines that form a complex web. This graphic serves as a background for the central text.  
**Almost 200 journalists  
Based in 65 countries**

*“Our aim is to bring journalists from different countries together in teams - eliminating rivalry and promoting **collaboration**. Together, we aim to be the **world’s best cross-border investigative team.**”*

[icij.org/about](http://icij.org/about)

# Exposure of hidden Secrets



Exposed the offshore holdings of 12 current and former world leaders.

And dealings of 128 more politicians and public officials around the world.

## The scale of the leak

Volume of data compared to previous leaks

**1,7 GB**

Cablegate/Wikileaks (2010)



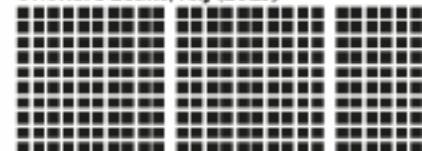
**≈ 2,6 TB**

Panama Papers/ICIJ (2016)

■ = 1GB

**260 GB**

Offshore Leaks/ICIJ (2013)



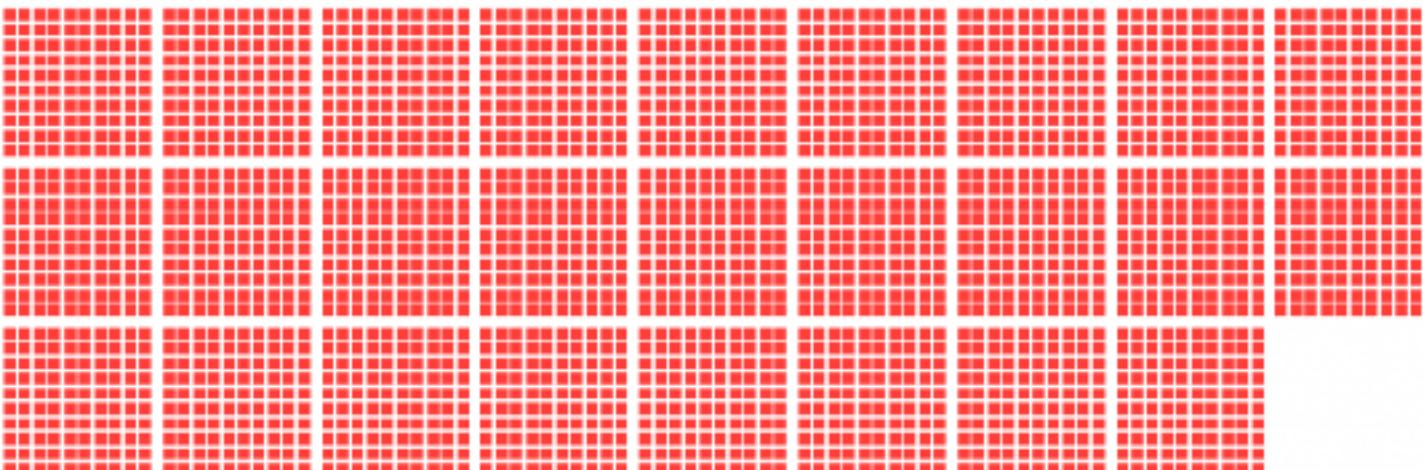
**4 GB**

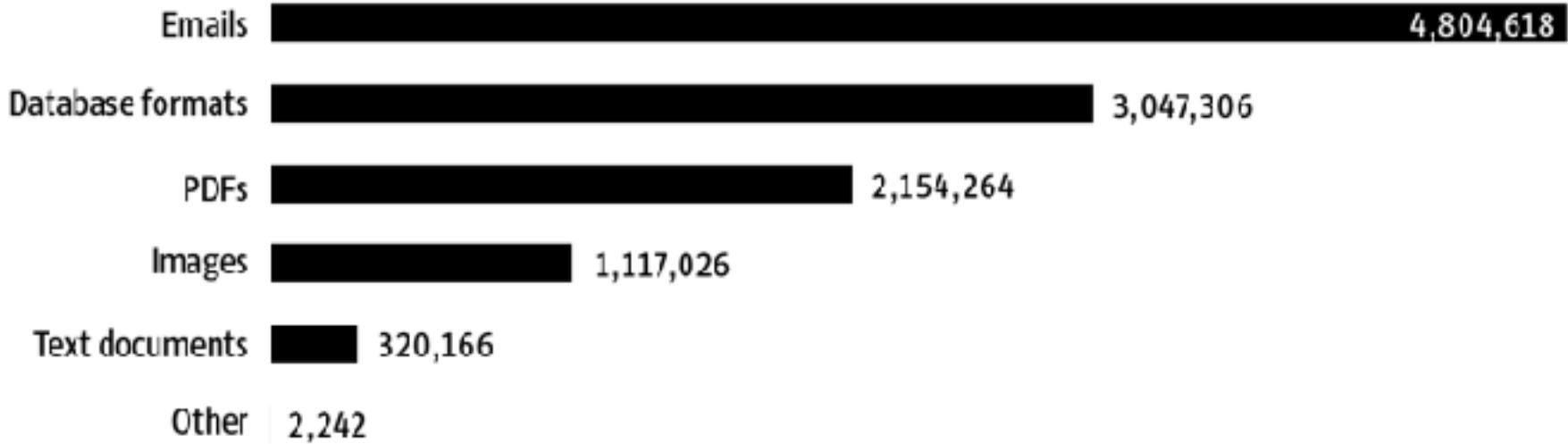
Luxemburg Leaks/ICIJ (2014)



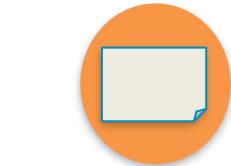
**3,3 GB**

Swiss Leaks/ICIJ (2015)





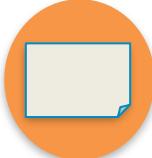
# Context is King



name:  
"Alice"  
last: „Smith“  
role:  
„Advisor“



name: "John"  
last: „Miller“  
role:  
„Negotiator“



name: "Jose"  
last: "Pereia"  
position:  
"Governor"

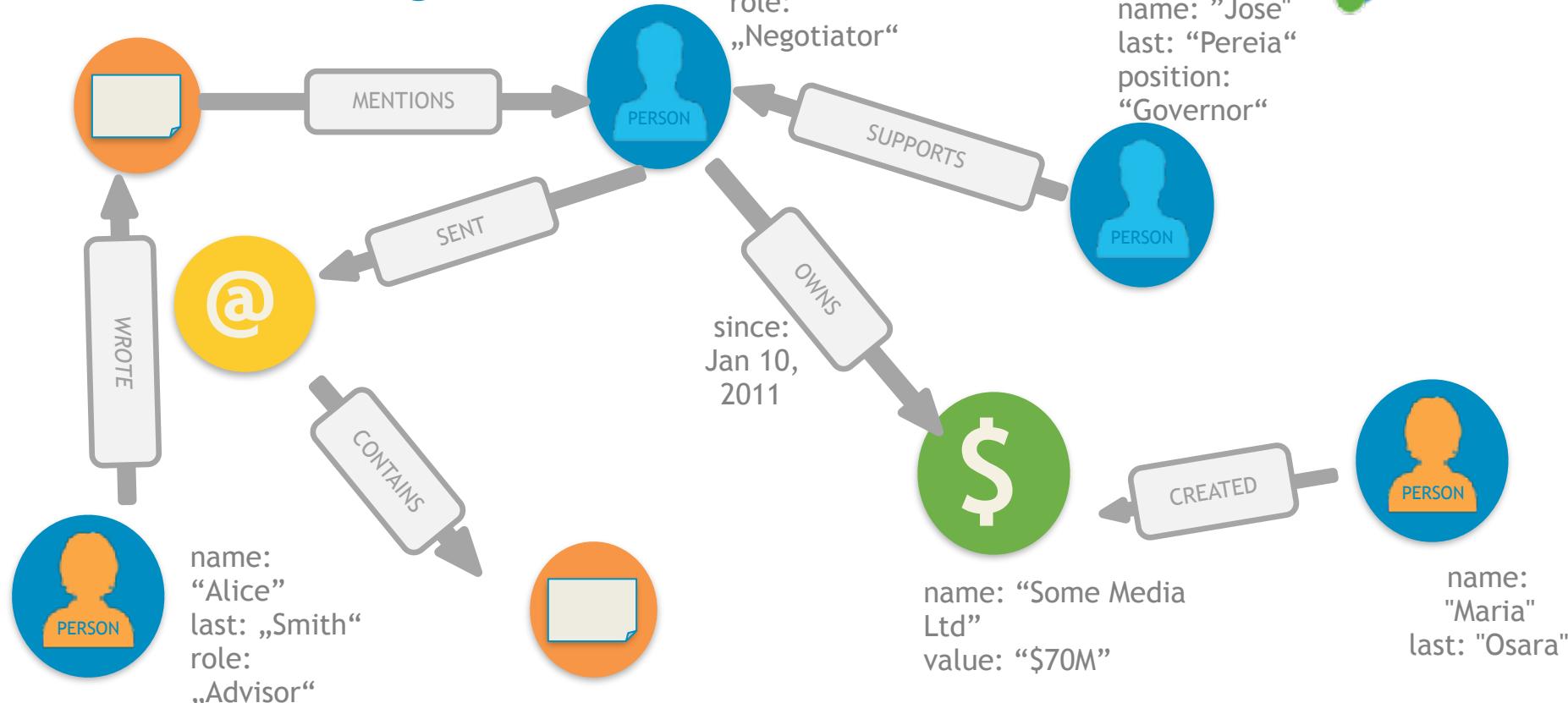


name: "Some Media  
Ltd"  
value: "\$70M"



name:  
"Maria"  
last: "Osara"

# Context is King



# We need a data model

Either based on our use cases & questions  
On the entities present in our meta-data and data.

## Meta Data Entities

- Document, Email, Contract, DB-Record
- Meta: Author, Date, Source, Keywords
- Conversation: Sender, Receiver, Topic
- Money Flows

## Actual Entities

- Person
- Representative (Officer)
- Address
- Client
- Company
- Account

# Data model - relationships

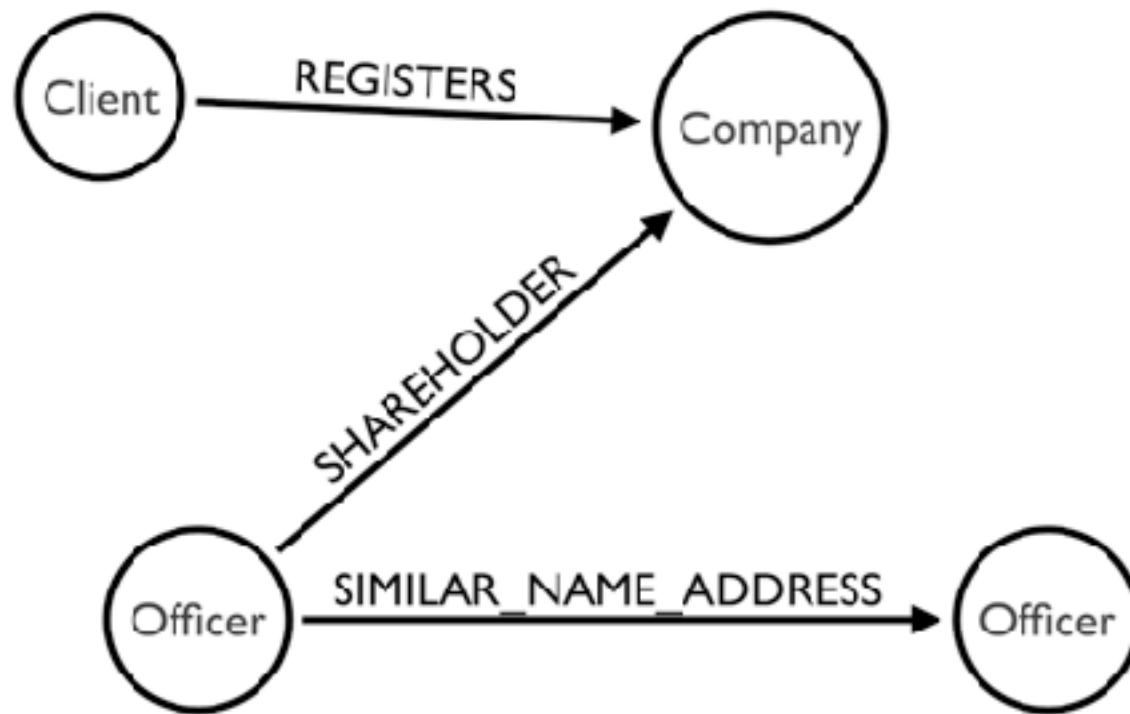
## Meta-Data

- sent, received, cc'ed
- mentioned, topic-of
- created, signed
- attached
- roles
- family relationships

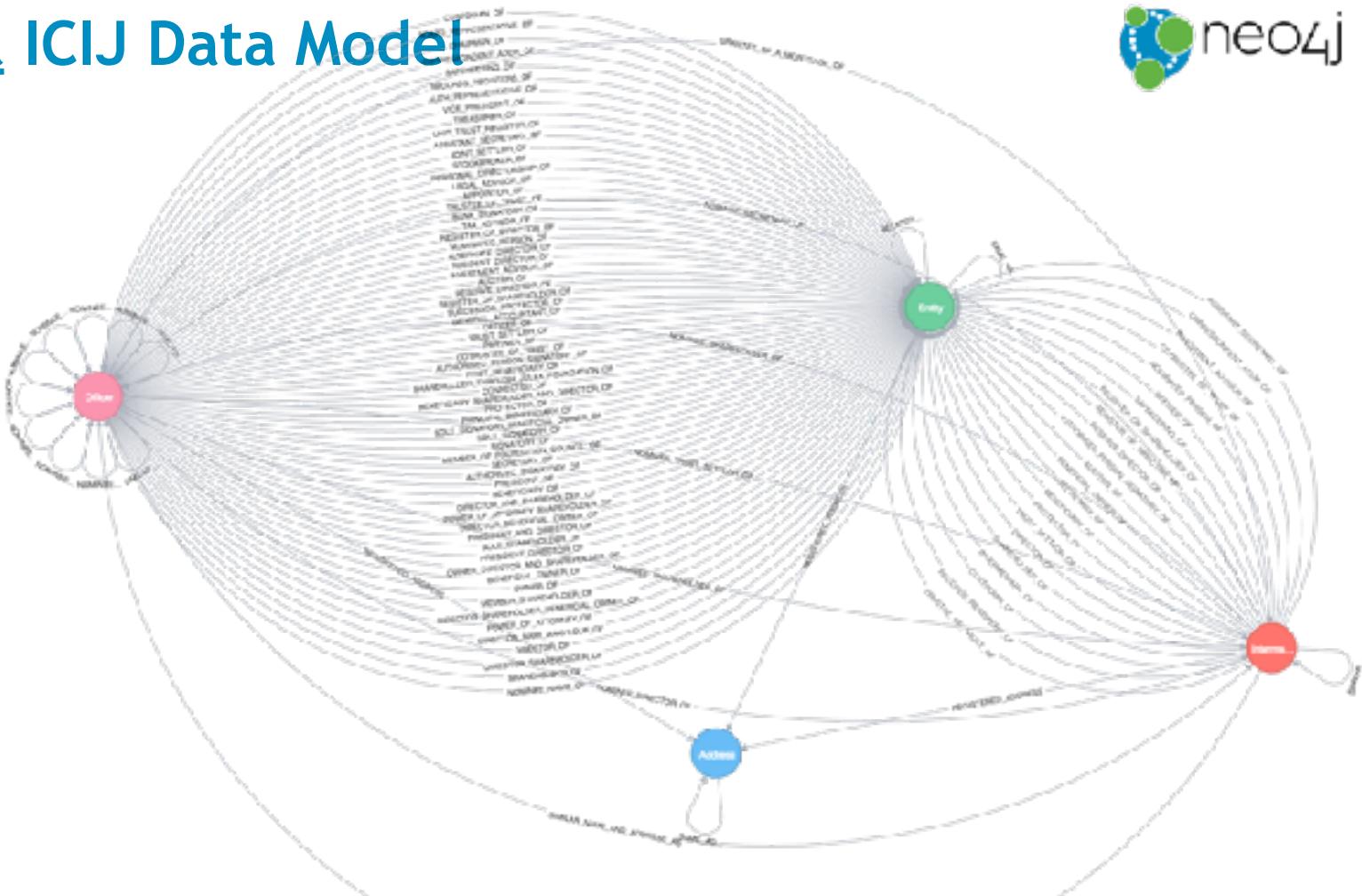
## Activities

- open account
- manage
- has shares
- registered address
- money flow

# The Basic ICIJ Data Model



# The Real ICIJ Data Model



# The Predictions

Country leaders Politicians/public officials Filters Region v Country v

A grid of 15 political portraits from various countries, arranged in three rows. The top row contains six portraits: President of Argentina, Former prime minister of Georgia, Iceland's prime minister, Ex-prime minister of Iraq, Former prime minister of Jordan, and Former prime minister of Qatar. The middle row contains six portraits: Former Emir of Qatar, King of Saudi Arabia, Former president of Sudan, UAE President, Abu Dhabi emir, Convicted former Ukraine prime minister, and President of Ukraine. The bottom row contains two portraits: Former prime minister of Australia and Australia's prime minister.

President of Argentina	Former prime minister of Georgia	Iceland's prime minister	Ex-prime minister of Iraq	Former prime minister of Jordan	Former prime minister of Qatar
Former Emir of Qatar	King of Saudi Arabia	Former president of Sudan	UAE President, Abu Dhabi emir	Convicted former Ukraine prime minister	President of Ukraine

Former prime minister of Australia Australia's prime minister

# Actually found because of Neo4j



# Paradise Papers: Commerce chief Wilbur Ross's links with sanctioned Russians

By Paradise Papers reporting team  
BBC Panorama

© 5 November 2017

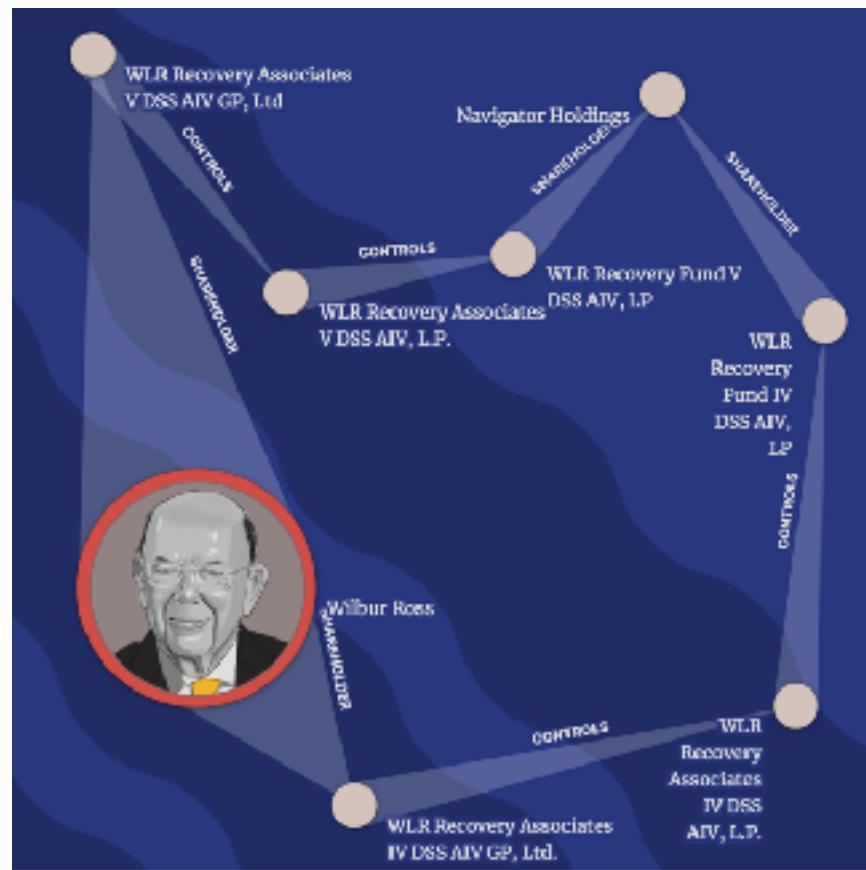
    Share



Wilbur Ross has played a key part in Donald Trump's business and political careers

A top member of Donald Trump's administration has business links with Russian allies of President Vladimir Putin who are under US sanctions, the Paradise Papers have revealed.

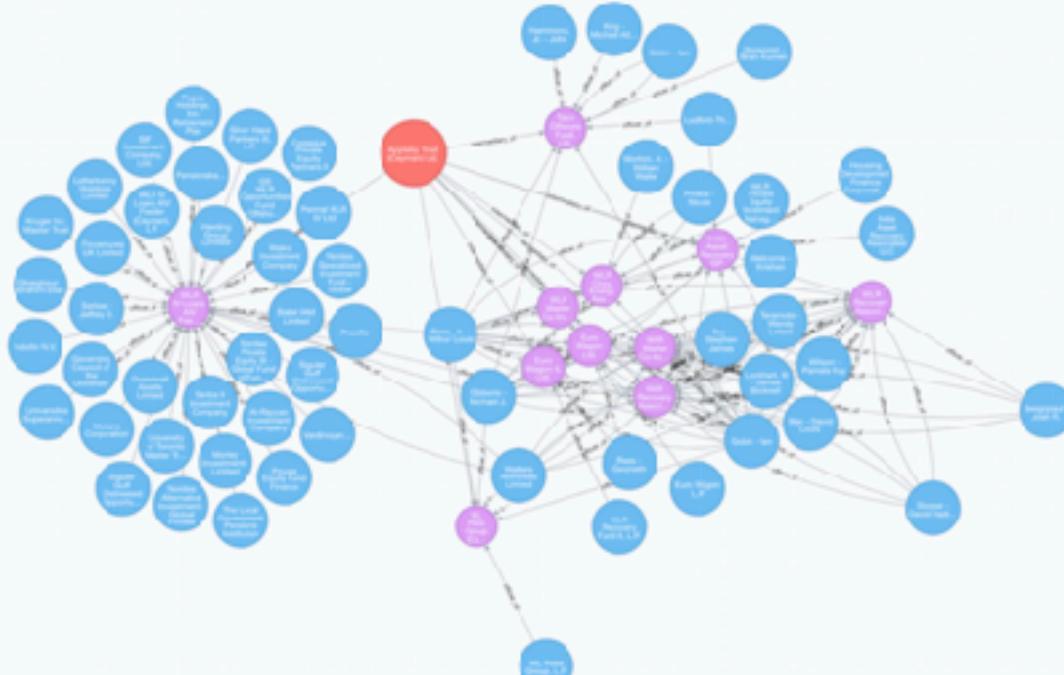
<http://www.bbc.com/news/world-us-canada-41876939>



<https://www.icij.org/investigations/paradise-papers/us-president-donald-trumps-influencers/>

# Bit more connected though

```
// Wilbur Ross's connections in the Paradise Papers
MATCH (o:Officer)-->(e:Entity)-[:intermediary_of]-(i:Intermediary)
WHERE o.name CONTAINS "Ross"
MATCH (e)--(o2:Officer)
RETURN *
```



# Make Sense Of Data

The screenshot shows the Neo4j Sandbox landing page with a grid of service cards:

- Input Service**: Jupyter service. Launch Service.
- Feedback Papers by ICN**: The Paradise Papers classification guide from the International Consortium members. Launch Service.
- Paradise Papers by ICN**: The Paradise Papers dataset, derived from customized documents from the International Consortium members (ICN). Launch Service.
- NeuroML Java**: Generate personalized real-time neuroinformatics - long distance research. Launch Service.
- Hosted AWS Lambda Functions**: Deploy analytical code snippets — no need to install or manage. Launch Service.
- Log4J**: Log4J Configuration generator. Launch Service.
- Markov Model**: Build Markov chains. Launch Service.
- Standardized DPLP**: Use our schema from Open Government. Launch Service.
- Twitter**: Explore trending Twitter topics, analyze sentiment, and extract entities. Launch Service.
- Open Data Catalog**: Search and browse over 1000 datasets from over 100 countries. Launch Service.
- GraphDB**: Store, query, and analyze graph data. Launch Service.

<https://neo4j.com/sandbox-v2/>

The screenshot shows the Neo4j Sandbox dashboard for the "Paradise Papers by ICN" service. The top bar includes "Logout" and navigation tabs: Get Started, Details, Data Mode, Code, Advanced, and a settings icon.

**Greetings William Lyon**  
Welcome to the Neo4j Sandbox. If you have any questions or problems, feel free to reach out to us at [davro@neo4j.com](mailto:davro@neo4j.com).

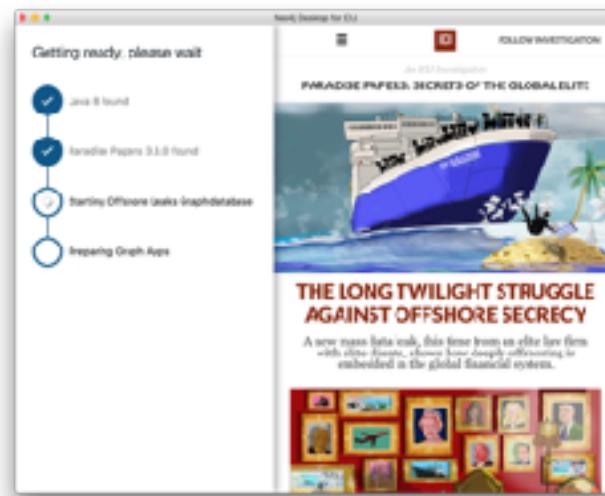
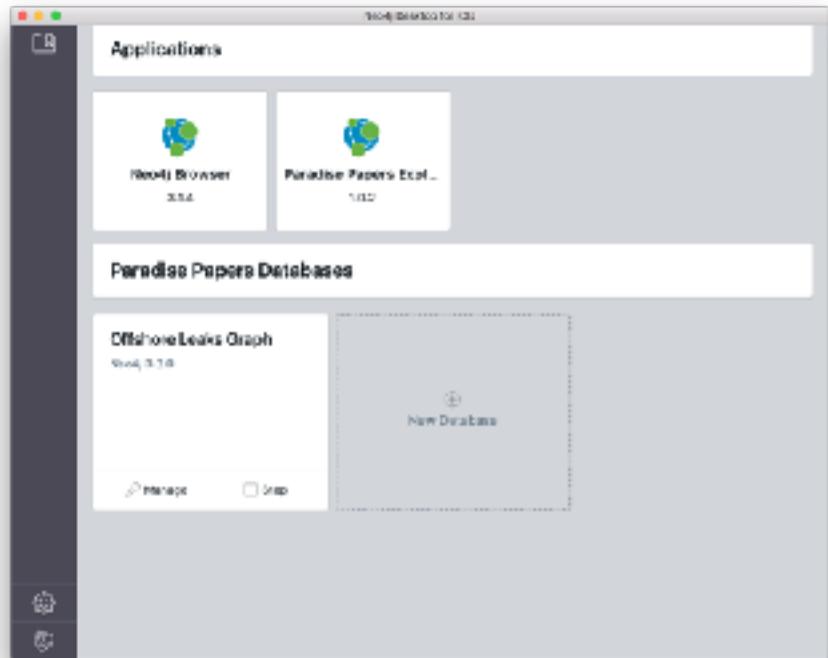
**Your Current Sandboxes**

**Paradise Papers by ICN**    Get Started    Details    Data Mode    Code    Advanced

**Neo4j Browser:** <https://192-168-1-201-205/neo4j/sandbox.com/>  
**Direct Neo4j HTTP:** <http://192.168.1.205:32875/browserv>  
**Username:** neo4j  
**Password:** cointinercas-licensed-john

**IP Address:** 192.168.1.205  
**HTTP Port:** 32875  
**Bolt Port:** 32874  
**Expires:** 2 days, 23 hours, 59 minutes

# Neo4j Desktop For ICIJ



<https://offshoreleaks.icij.org/pages/database>

<https://www.eventbrite.com/e/introduction-to-neo4j-san-diego-tickets-56209010664>

The image shows a screenshot of an Eventbrite event page. At the top, there's a large photo of three people smiling, overlaid with a network graph visualization. To the right of the photo, the date "MAR 19" is displayed. Below the date, the event title "Introduction to Neo4j - San Diego" is shown, followed by the organizer "by Neo4j". The price "\$149 - \$299" is listed next. A prominent green button at the bottom right says "Tickets".

**Description**

**Introduction to Neo4j**  
Duration: 8-hours/1 day  
Skill Level: Beginner

Audience for this course: Developers, Architects, Administrators, Data Scientists, Data Analysts

**Course Description**

This course introduces you to what a graph database is and how

**Date And Time**

Tue, March 19, 2019  
9:30 AM – 5:00 PM PST  
[Add to Calendar](#)

**Location**

San Diego Training & Conference Center  
450 B St



<https://www.eventbrite.com/e/graphtour-san-francisco-ca-tickets-58600670182>

eventbrite

Search for events

Browse Events

Help

Sign In

Create Event



MAY  
08

GraphTour - San Francisco, CA

by Neo4j

Free

Register



#### Description

Neo4j is hitting the road to bring a full day of content-rich sessions on how graph databases are revolutionising the modern enterprise.

Meet our experts to hear first-hand about the advantages of Neo4j's native Graph Platform, which offers not just the Neo4j database, but also Analytics, Data Import and Transformation, Visualization, and Discovery.

#### Date And Time

Thu, May 9, 2019

9:00 AM – 5:00 PM PDT

[Add to Calendar](#)

Innovation



<https://maxdemarzi.com>

# MAX DE MARZI

Graphs with Neo4j

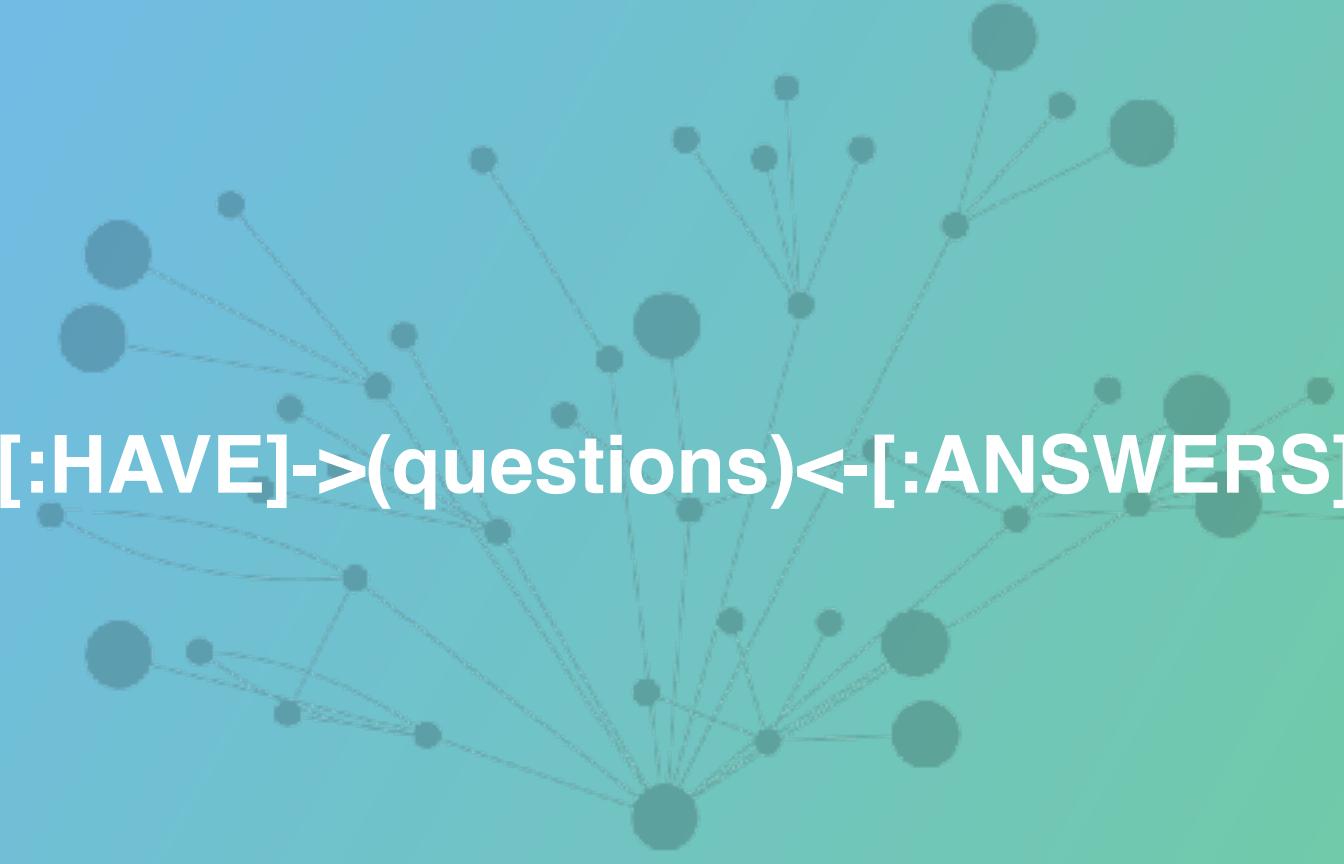
[VIDEOS](#) [SERVICES](#) [CONTACT](#) [ABOUT](#) [HOME](#)

MAR 01 2019  
[LEAVE A COMMENT](#)

JAVA

NETWORK ROUTING IN NEO4J



A network graph visualization consisting of numerous small, semi-transparent teal-colored circles of varying sizes scattered across a light blue background. These nodes are interconnected by a dense web of thin, greyish-blue lines, forming a complex web-like structure.

(you)-[:HAVE]->(questions)<-[ :ANSWERS ]-(neo)