GRAPHISTRY

Scaling graph visualization: How & Why

Graph + AI, 2021

Leo Meyerovich, CEO
@LMeyerov

info@graphistry.com

GRAPHISTRY

100X VISUAL INVESTIGATIONS

Users Security, fraud, user analytics, genomics, …

Tech Graph, viz, GPUs, automation

info@graphistry.com

Unlocking graph for more people

TigerGraph

GRAPHISTRY

- Why visualize graphs
- Bringing to everyone with no-code & low-code
- Tech: Scale visual insights with smart UI & GPUs

DEMO Project Domino: Citizen science against medical misinformation

- Help the good
- Fight the bad

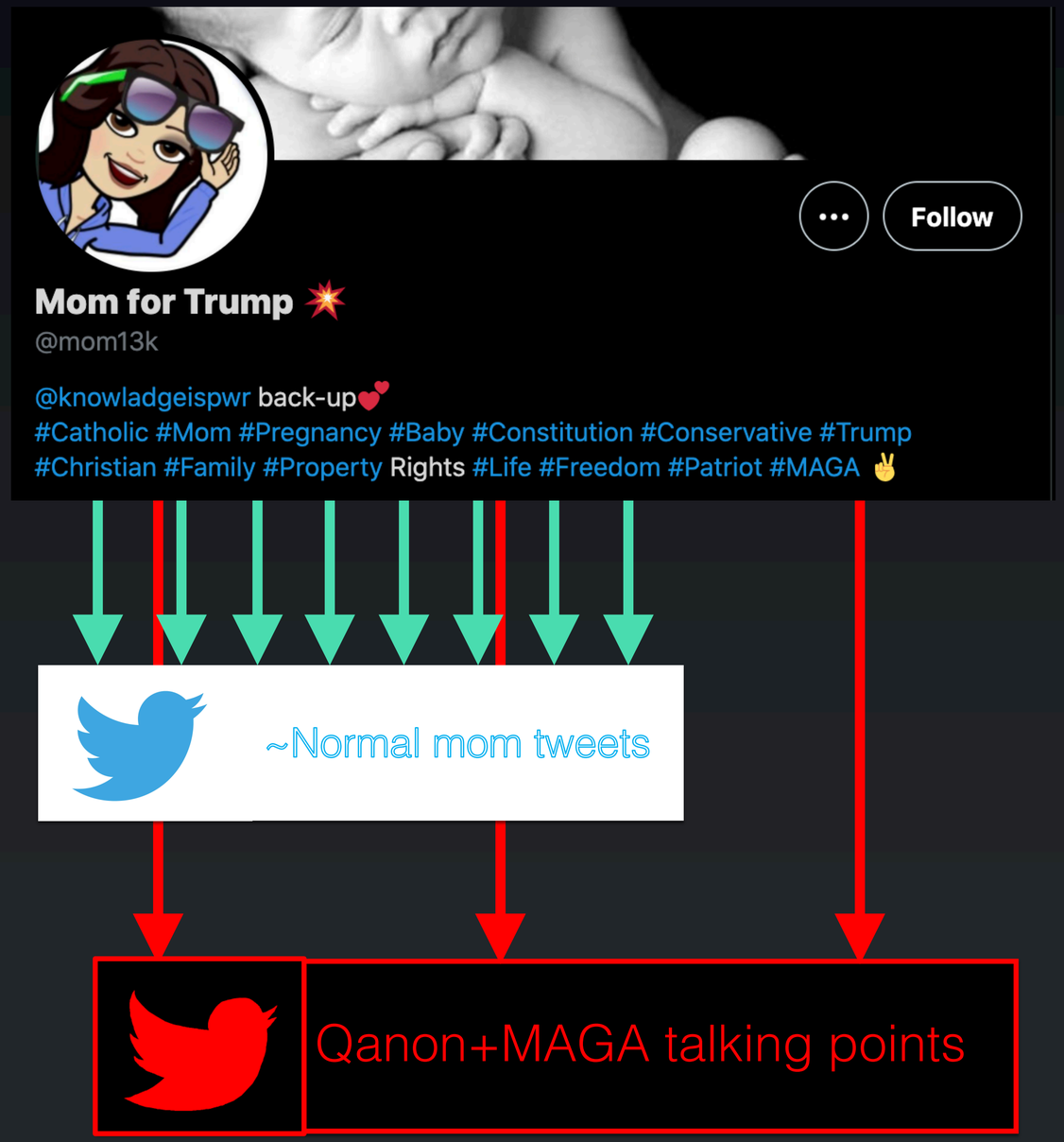
See DefCon AIV keynote

PROJECT DOMINO Citizen data science against misinformation

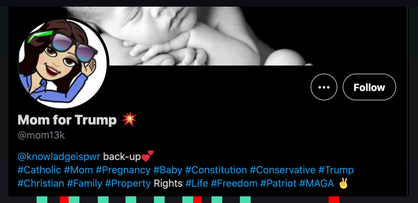
2021 HBO documentary unraveling qanon leadership & Trump ties

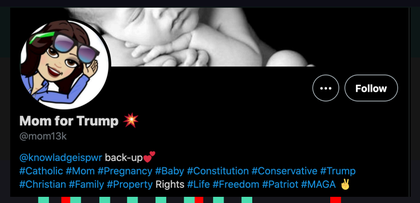
Today: The investigation behind the big reveal in episodes 5 + 6

SOURCES
Cody Webb (Project Domino)
Tom Porter (Business Insider)
Cullen Hoback & Adam McKay (HBO)
Testimonials of caught instigators

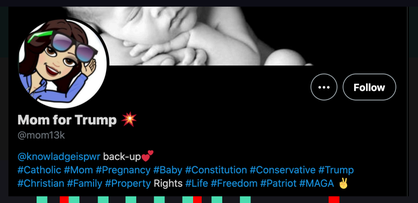

Follow
Mom for Trump
@mom13k
@knowladgeispwr back-up
#Catholic #Mom #Pregnancy #Baby #Constitution #Conservative #Trump
#Christian #Family #Property Rights #Life #Freedom #Patriot #MAGA
~Normal mom tweets
Qanon+MAGA talking points

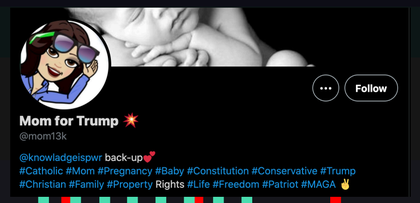
Follow
Mom for Trump
@mom13k
@knowladgeispwr back-up
#Catholic #Mom #Pregnancy #Baby #Constitution #Conservative #Trump
#Christian #Family #Property Rights #Life #Freedom #Patriot #MAGA
~Normal mom tweets
Qanon+MAGA talking points

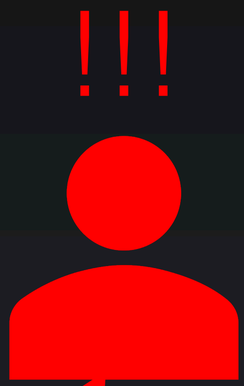

Follow
Mom for Trump
@mom13k
@knowladgeispwr back-up
#Catholic #Mom #Pregnancy #Baby #Constitution #Conservative #Trump
#Christian #Family #Property Rights #Life #Freedom #Patriot #MAGA
Normal mom tweets
Qanon+MAGA talking points


Follow
Mom for Trump
@mom13k
@knowladgeispwr back-up
#Catholic #Mom #Pregnancy #Baby #Constitution #Conservative #Trump
#Christian #Family #Property Rights #Life #Freedom #Patriot #MAGA
Normal mom tweets
Qanon+MAGA talking points

???







~Normal mom tweets

Qanon+MAGA talking points

TWEET METADATA

DOMAIN REGISTRAR

Power10App.com

Power10App.com

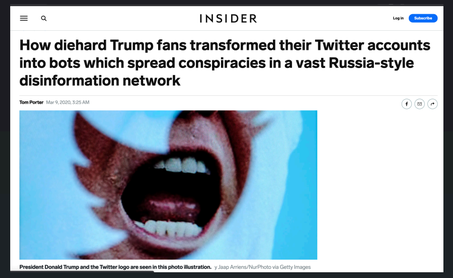
JASON SULLIVAN
Mueller witness

Social media manager:
Tools, campaigns, one-offs

Trump campaign

ROGER STONE
convicted 2019

Election crimes
+ Proud Boys sponsor



INSIDER

How diehard Trump fans transformed their Twitter accounts into bots which spread conspiracies in a vast Russia-style disinformation network

Tom Porter Mar 9, 2020, 3:25 AM

President Donald Trump and the Twitter logo are seen in this photo illustration. | Jaap Arriens/NurPhoto via Getty Images

2018: Investigation

2020: BI article

Power10App.com

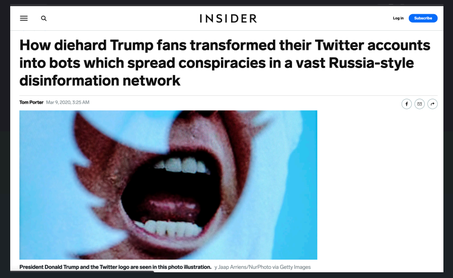
JASON SULLIVAN
Mueller witness

Social media manager:
Tools, campaigns, one-offs

Trump campaign

ROGER STONE
convicted 2019

Election crimes
+ Proud Boys sponsor



HBO

2018: Investigation

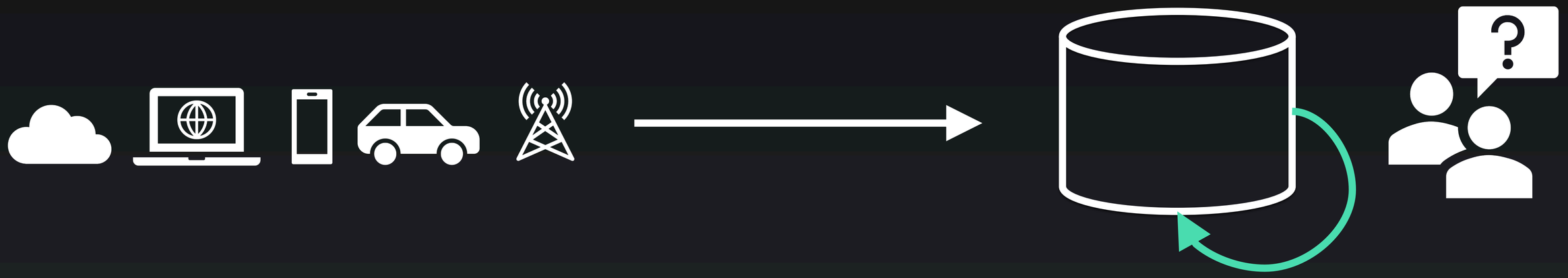
2020: BI article

2021: Qanon reveal

8Chan + QAnon

RON WATKINS

Public admin
admitted power10 runner,
admitted Q (takeover)

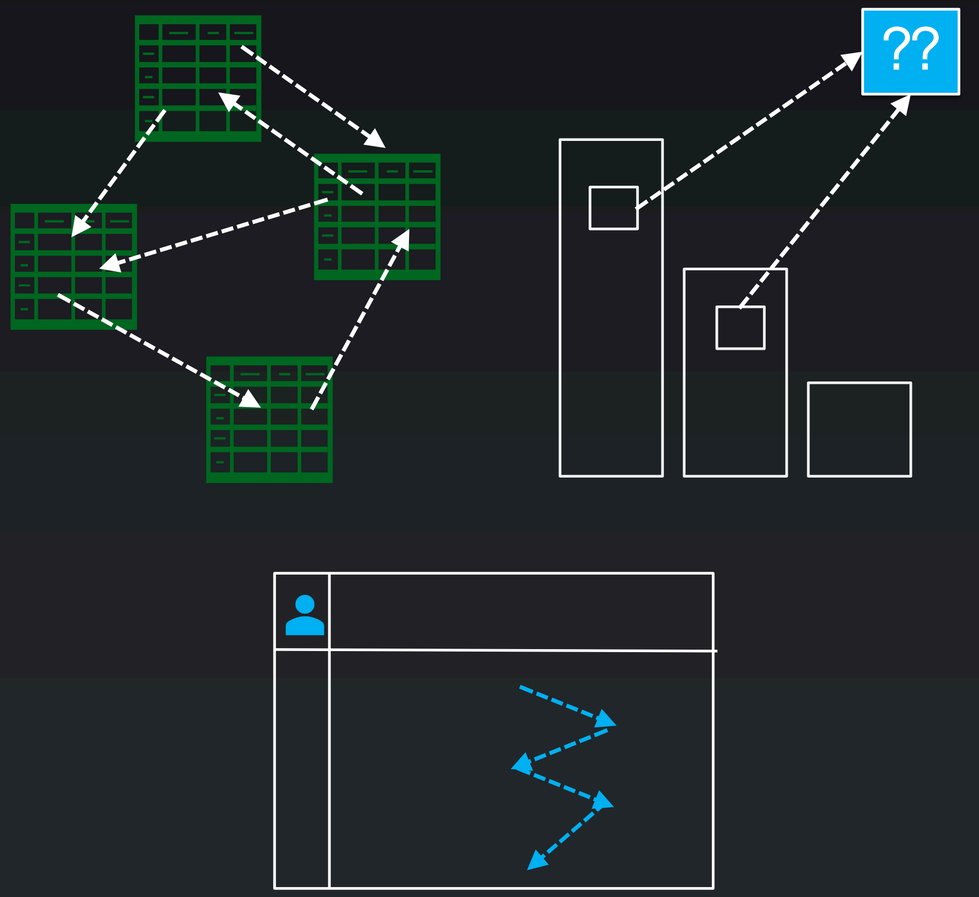


OPERATIONAL DATA IS HARD
Sales, security, fraud, IT, genomics, …

We now collect a ton of data from everything

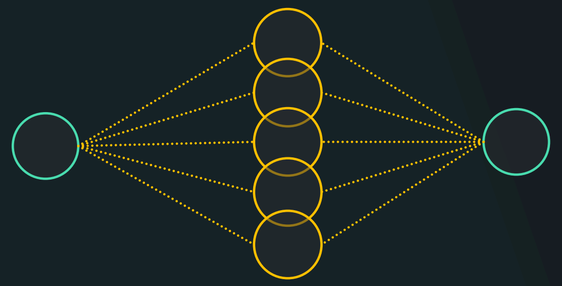
What are the relationships around the entities & events?
Scope, path, root cause, behavior, patterns & outliers, …

Hard to see relationships
with tables and tiny static charts

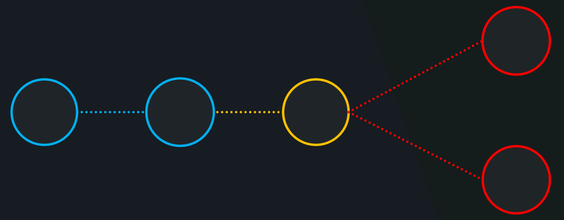


Graph brings visual understanding to relationships & AI

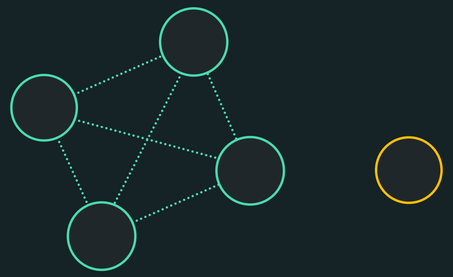
Scope



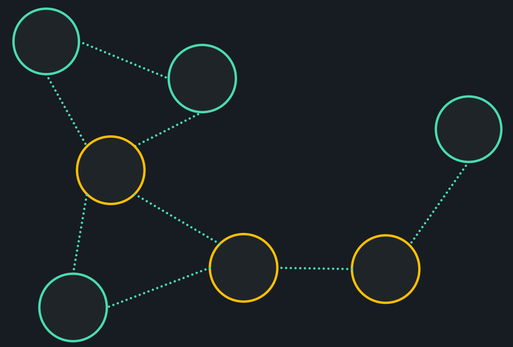
Progression & Behavior



Patterns & Outliers



Score influence, surprise, …

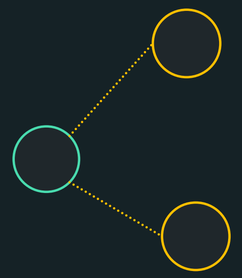


UI for AI:
Explore inferred scores, relationships, clusters, ..

GRAPHISTRY

© 2018 Graphistry, Inc. All rights reserved. Confidential and proprietary information. Do not distribute.

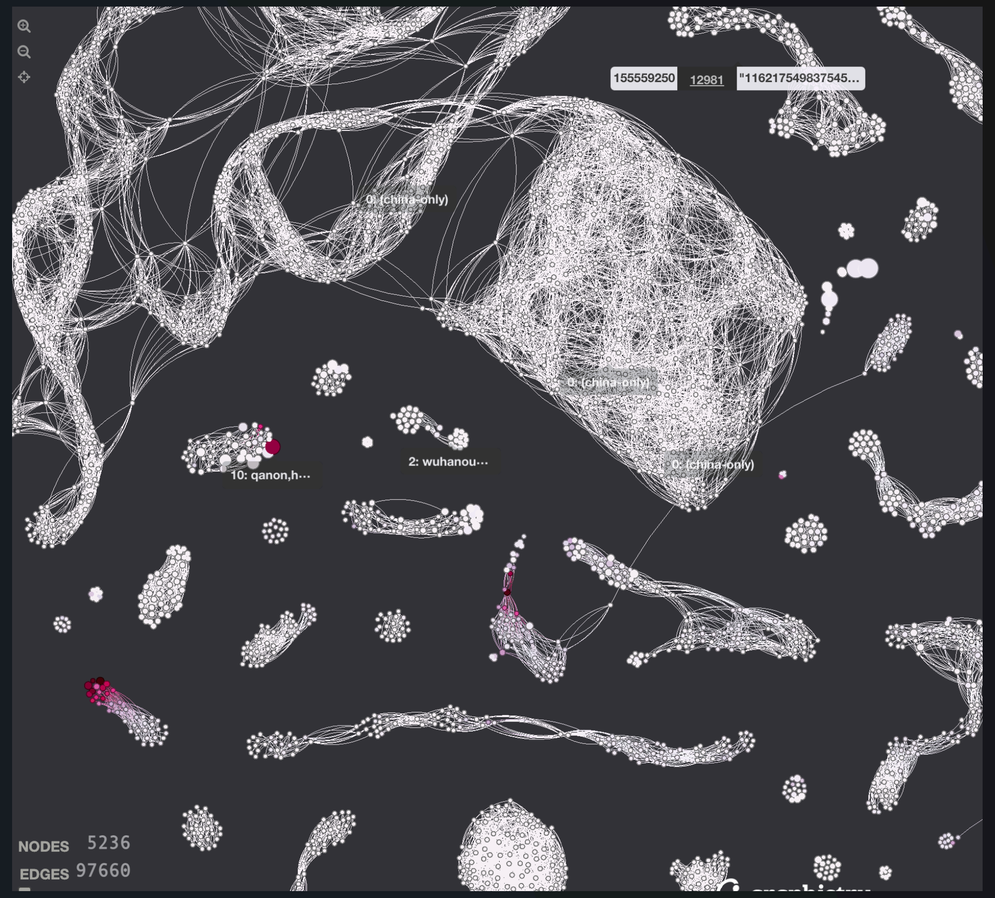
SMALL GRAPH TOOLS
An ant's eye view for local insights



patterns
outliers
entities
groups
behaviors
…

interaction loop

SCALABLE TOOLS
Eagle's perspective for operational data



Data from operational systems:
Incident analysis, fraud rings, social networks, demographic analysis, DNA, …

Designing for scale

Show a lot of data

GPU-scaled rendering
ML+stats driven design
Information design
Variety: Timebars, histograms, ..

Fast question/answer loop

No-code interactions
GPU-accelerated interactions
Summaries: Histograms, …
Color by dimension
Filter (cross-linked)

GPU accelerate what is possible in 100ms: Aim for GB/s and TFLOPS

GPU rendering engine

GPU analytics engine

Live stream

Database

1 GB/s

Big & fast data pushdown

Storage array

1 GPU
1-12 GB RAM

1MB/s – 1GB/s

1-8 GPUs per server
12-640GB GPU RAM

90 GB/s
1 TB – 1 PB

20ms

1-40ms

30ms - 10s

GRAPHISTRY

APACHE ARROW

NVIDIA

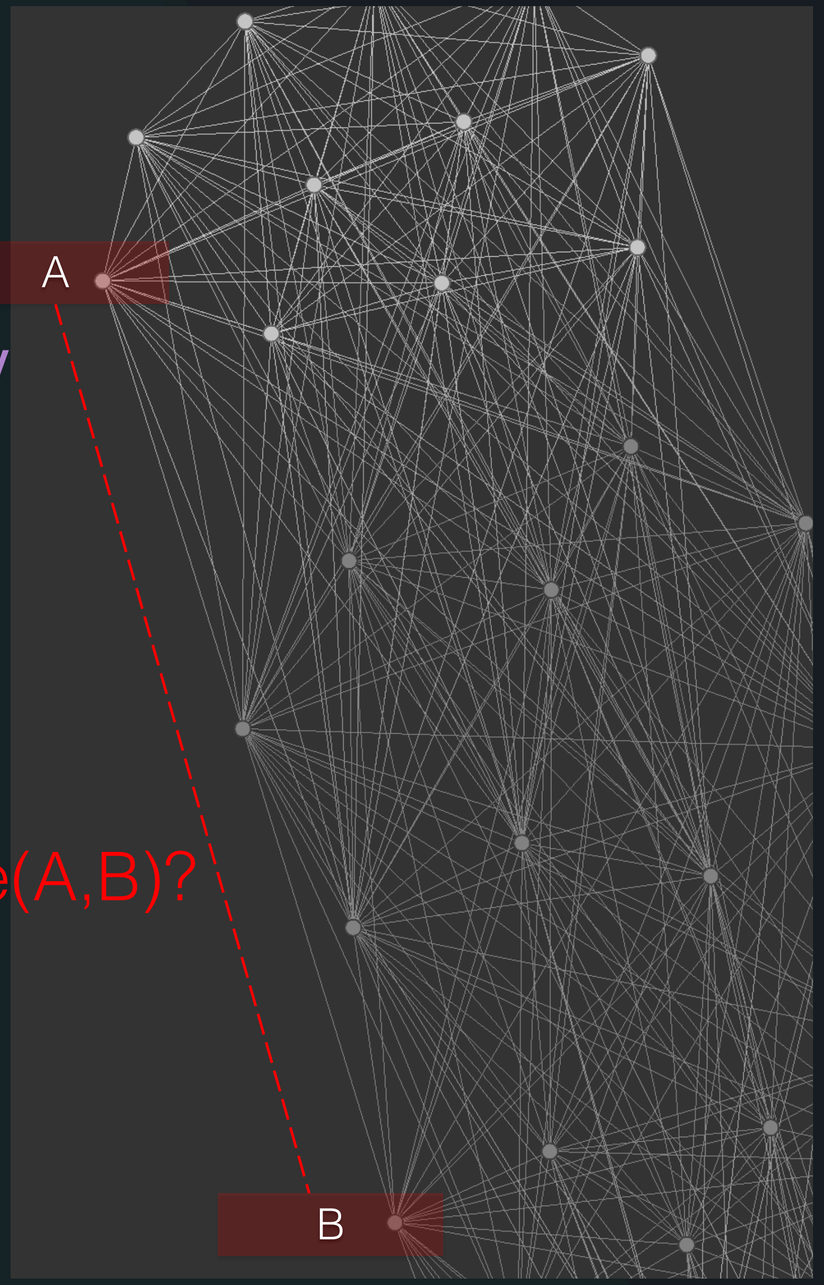
RAPIDS

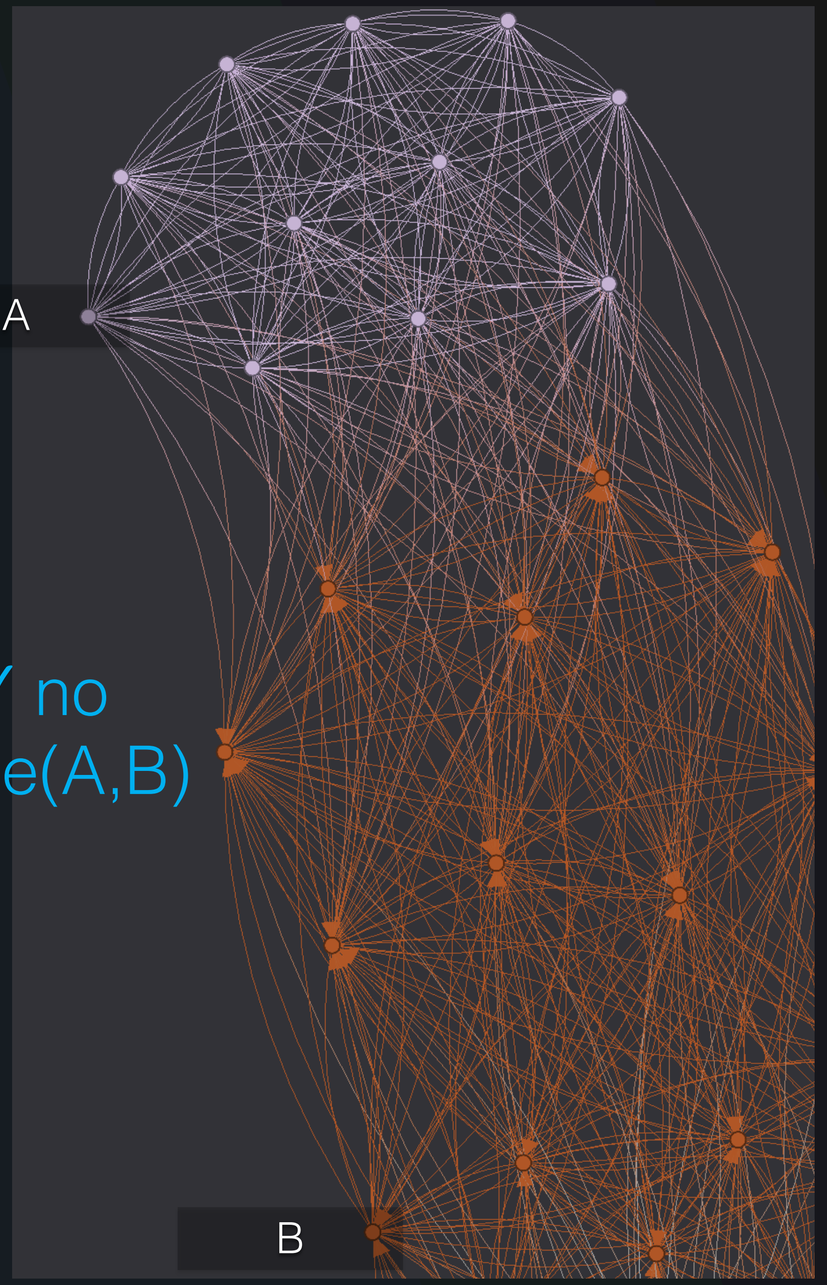
Ex: Scalable design for colors & curvature

Color reveals community for node, edge

Curvature reveals directed flows and disambiguates lines

Repeat for:
Layout, sizes, points of interest, zoom levels, picking, …

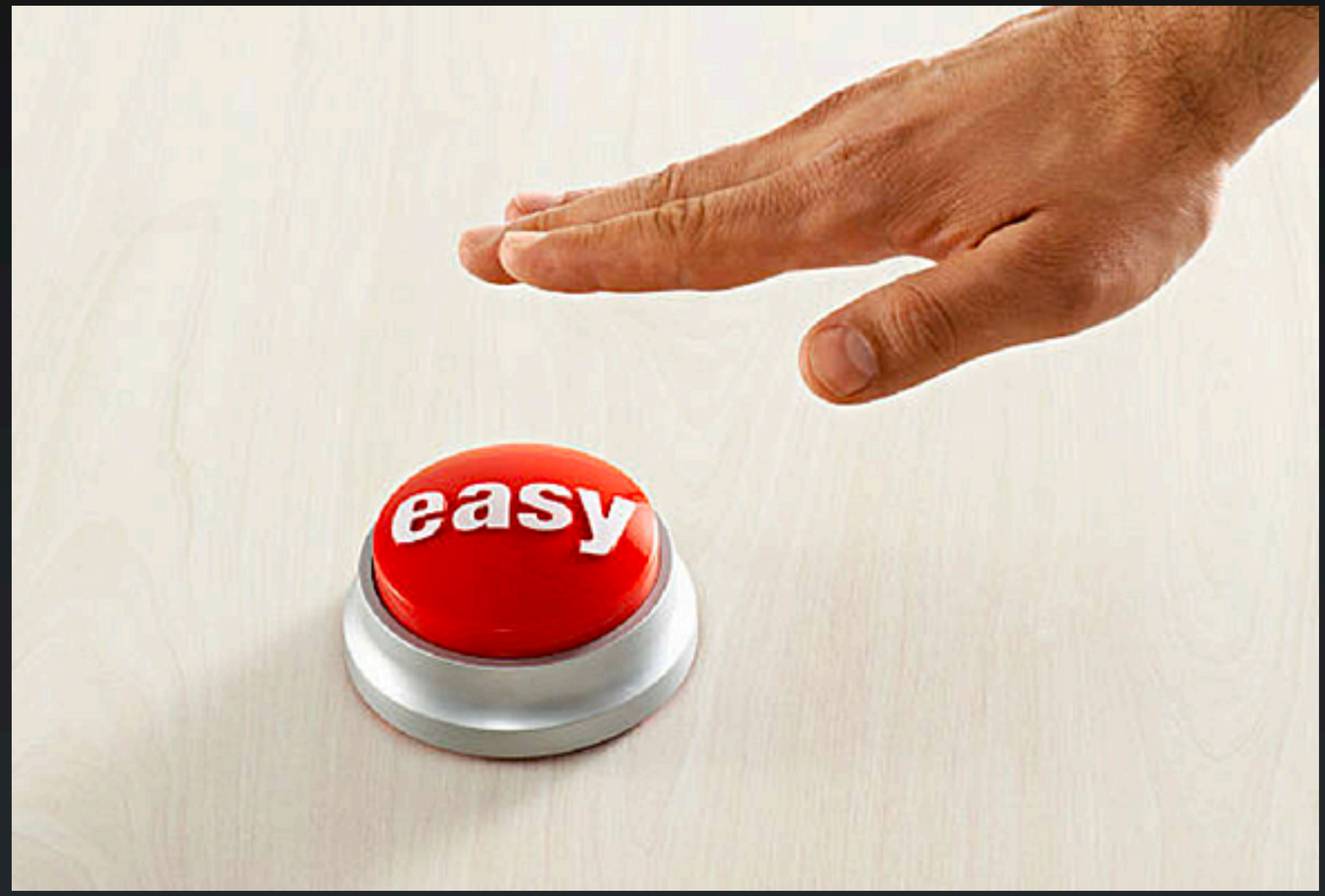




PROBLEM
85% of data science projects fail

- Gartner, 2020

Most people are busy so we need Easy Buttons



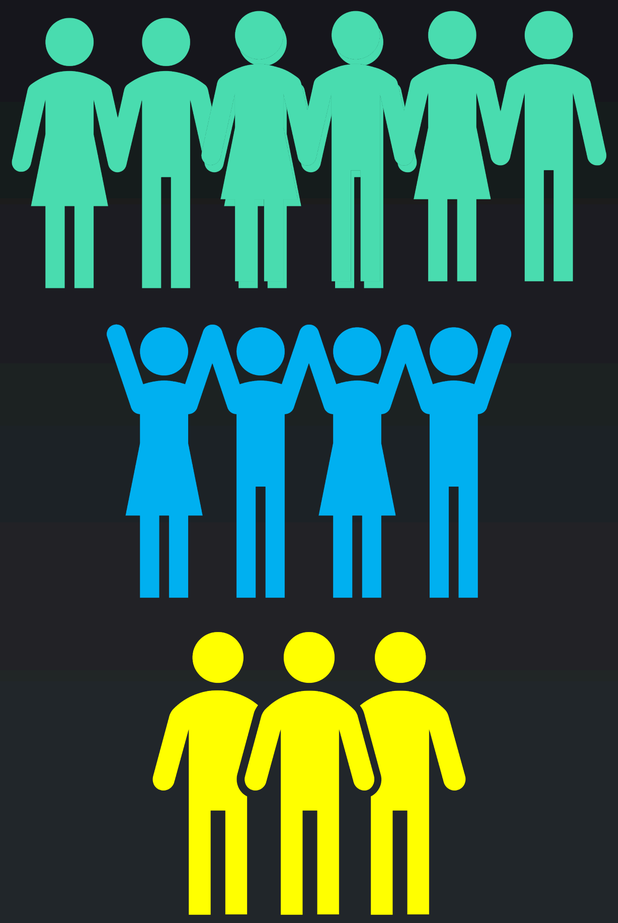
Graph is useful for coders + non-coders

data scientist, data engineer, dev, …

scientist, analyst, investigator, …

sales, marketing, trading, ...

Multiply the power, reach, & ROI of graph projects by rethinking coding as UX feature, not a gate keeper



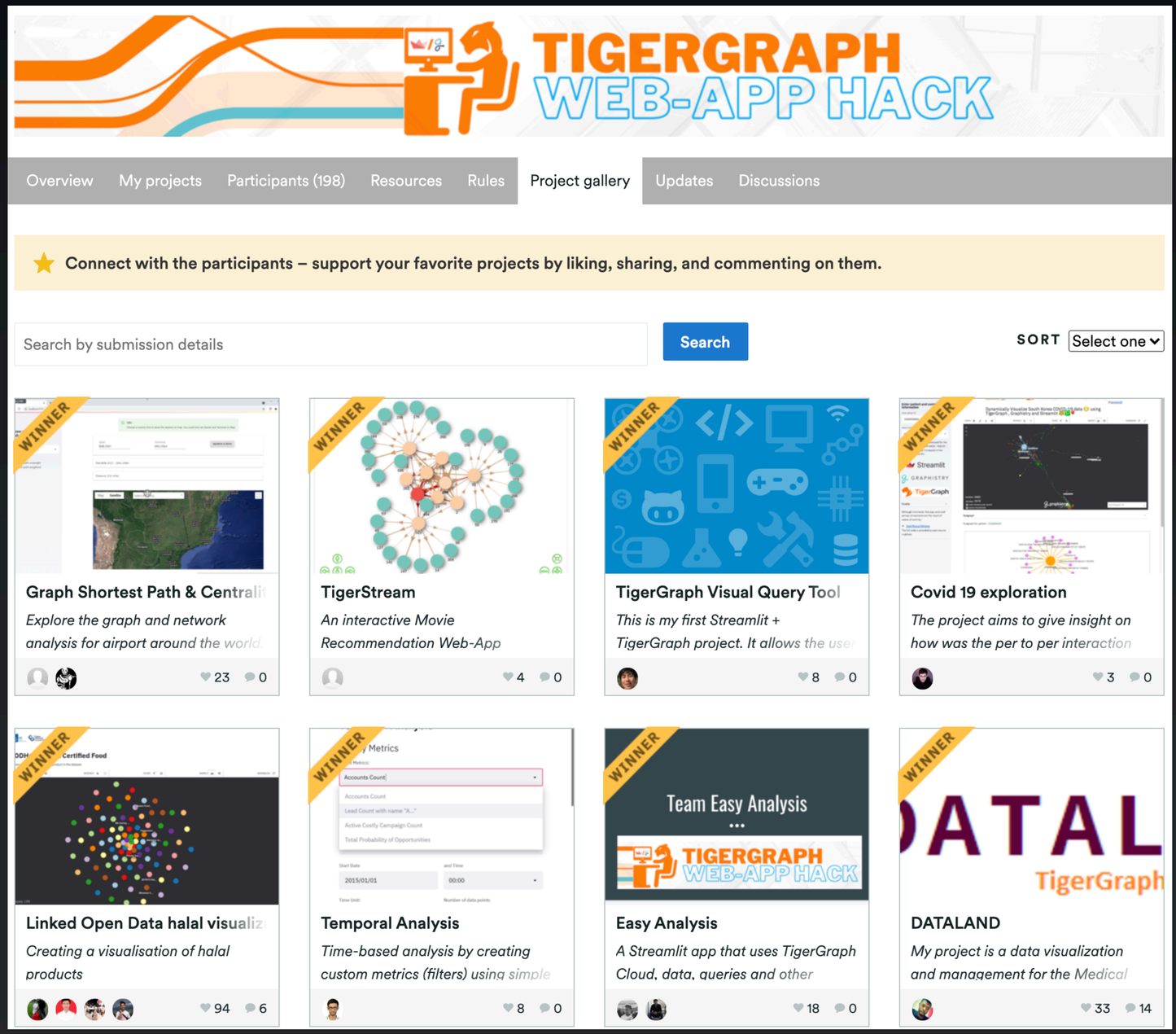
No-code

↑

Low-code create dashboards + automations

↑

Embedding APIs supercharge existing tools



Streamlit

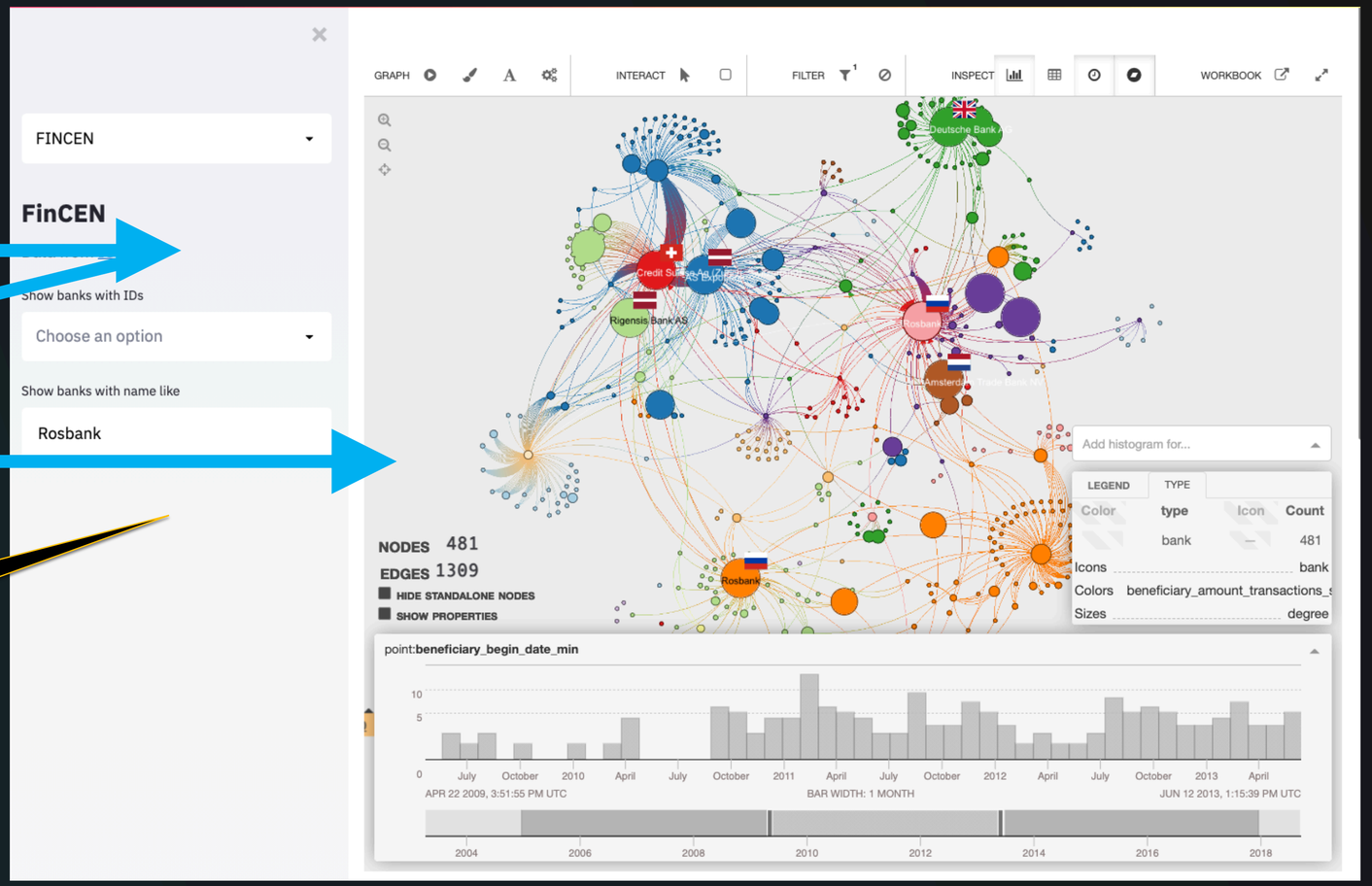
GRAPHISTRY

TigerGraph

graph-app-kit pipeline template

Most apps are copy/edit of: Form controls -> run queries & filter -> viz

```
def run_all():

    custom_css()

    try:

        # Render sidebar and get current settings
        sidebar_filters = sidebar_area()

        # Compute filter pipeline (with auto-caching based on fi
        # Selective mark these as URL params as wel
        filter_pipeline_result = run_filters(**si    ar_filters)

        # Render mai  viz ar    based on computed filter pipeline
        main_area(    ebar_filters,  **filter_pipeline_result)

    except Exception as exn:
        st.write('Error loading dashboard')
        st.write(exn)
```



Low-code most apps:
Make a few big input controls, and Graphistry UI for rest

github.com/graphistry/graph-app-kit

Quicklaunch, Tigergraph tutorials, Slack, & more

Low-code II:
Can we explore & automate without Python?

Low-code III:
Data science & GPUs

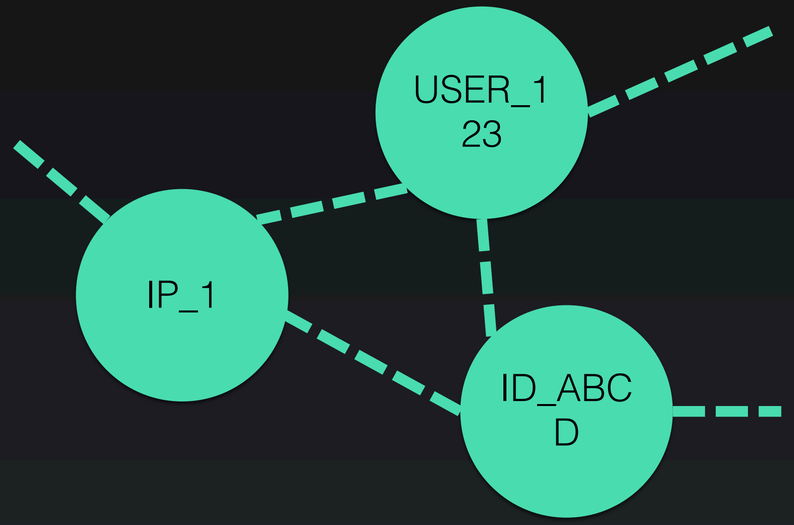
PyGraphistry: Notebook graph one-liners + team notebook sharing

github.com/graphistry/pygraphistry

- TigerGaph (notebook demo)

```
g = graphistry.tigergraph(protocol='https', ...)
g2 = g.gsql("...", {'edges': '@@eList'})
g2.plot()
print('# edges', len(g2._edges))
```

```
g.endpoint('my_fn', {'arg': 'val'}, {'edges': '@@eList'}).plot()
```



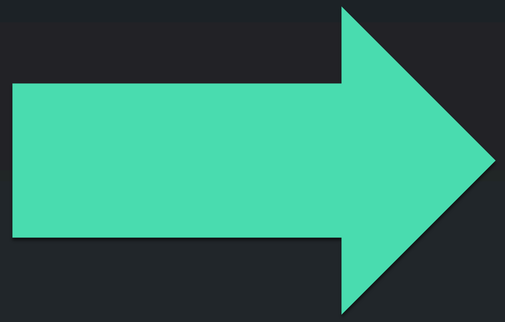

```
table_rows = pd.read_csv('honeypot.csv')
graphistry.hypergraph(table_rows, ['attackerIP', 'victimIP', 'victimPort'
```

1 GPU Speedups

```
time_max =
  cudf
    .read_parquet('/file.parquet')
    ['_time']
    .max()
```

In-memory GPU is amazing & ~easy insights

1 line change
(same API!)



- Graphistry: ~100X over non-GPU (crash -> fast)
- cuDF rollup: ~100X over pandas (20s -> 0.4s)
- In-memory: ~20X over cold

Takeaway

1. Unlock operational data insights via graph visual analytics
2. GPU era for graph is incredible
3. Multiply ROI by no-code & low-code views of graph

Thank you!

Try for yourself:

1. Graphistry.com – try a CSV, free API key, launch private VM
2. github.com/graphistry/pygraphistry & github.com/graphistry/graph-app-kit
3. Chat: Slack, ZenDesk, email, …

Tigergraph, no-code, SaaS, …: Contact for design partnering