



Verum

How Skynet Started as a Context Graph

Agenda

Me

- Gabriel Bassett
- Infosec graph guy
- Verizon DBIR contributing author
- Other stuff

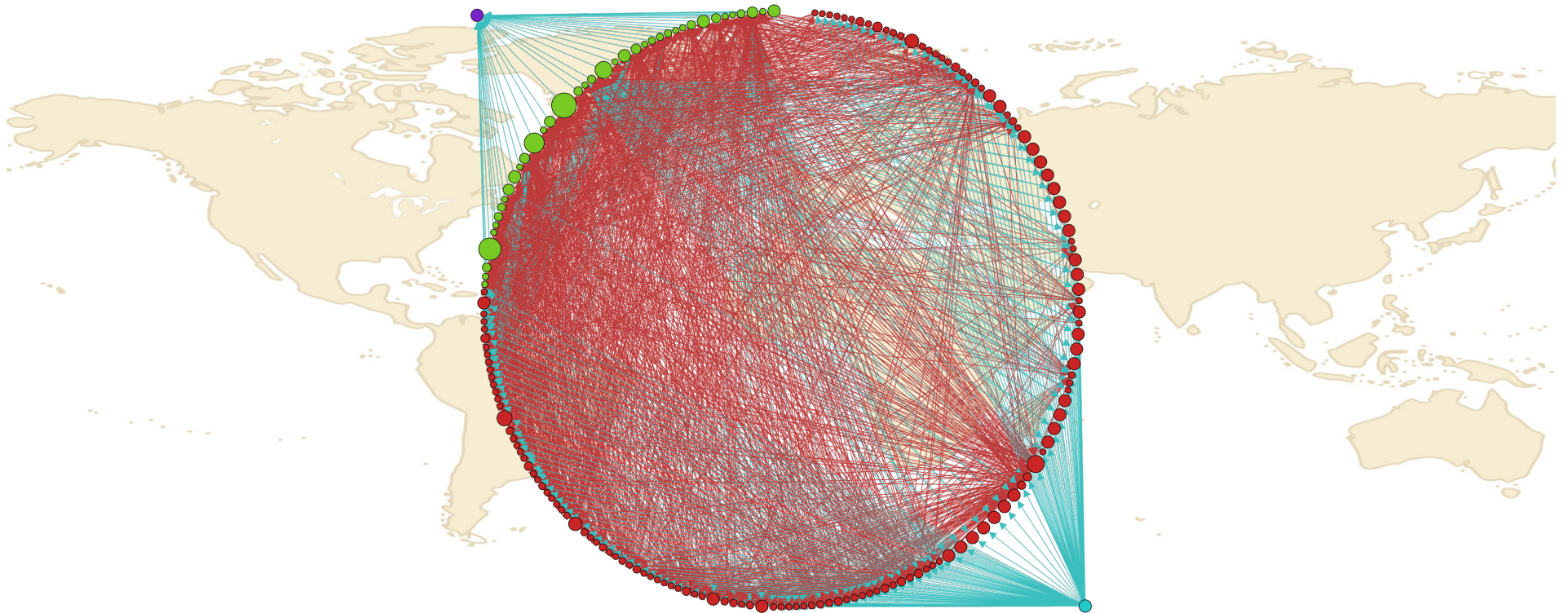
Agenda

- Understanding the problem
- Why threat intel sucks
- Graph theory 101
- My solution - Verum
 - Schema
 - Implementation
 - Teaching
 - Learning
 - Thinking
 - Communicating

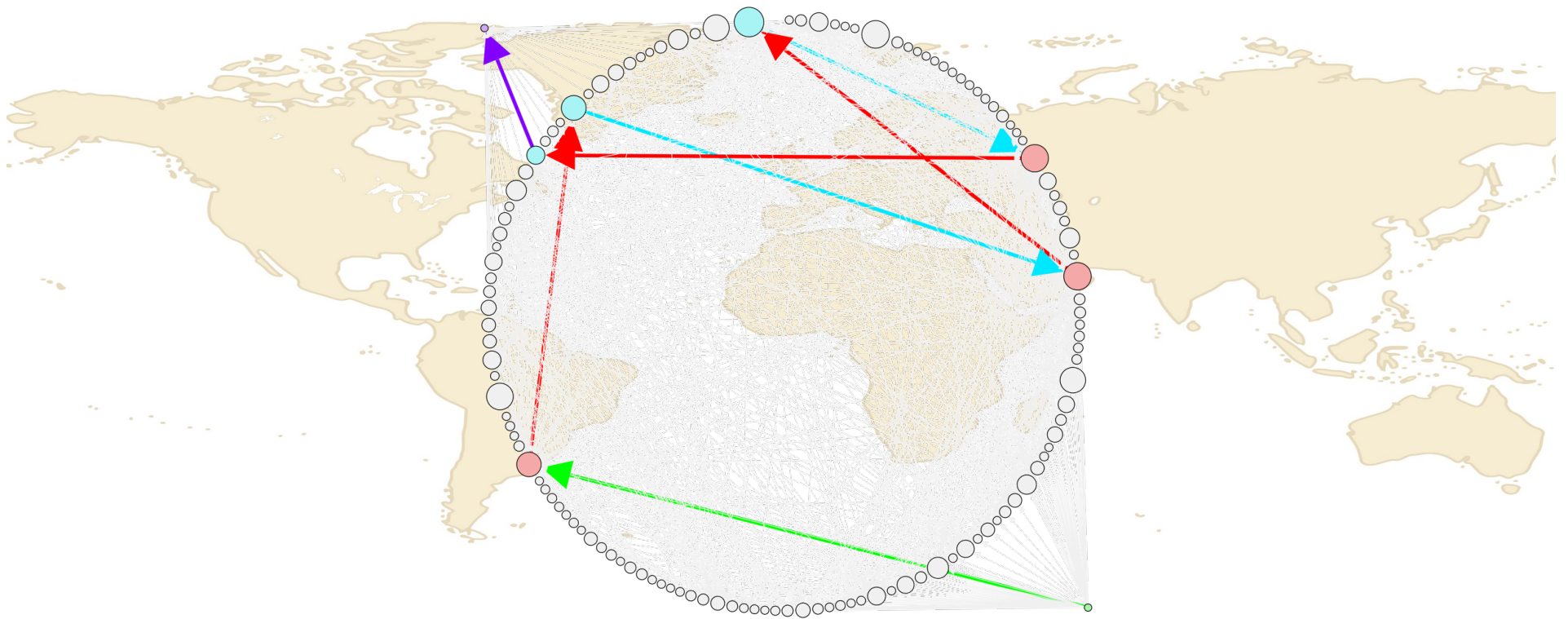


Understanding the Problem

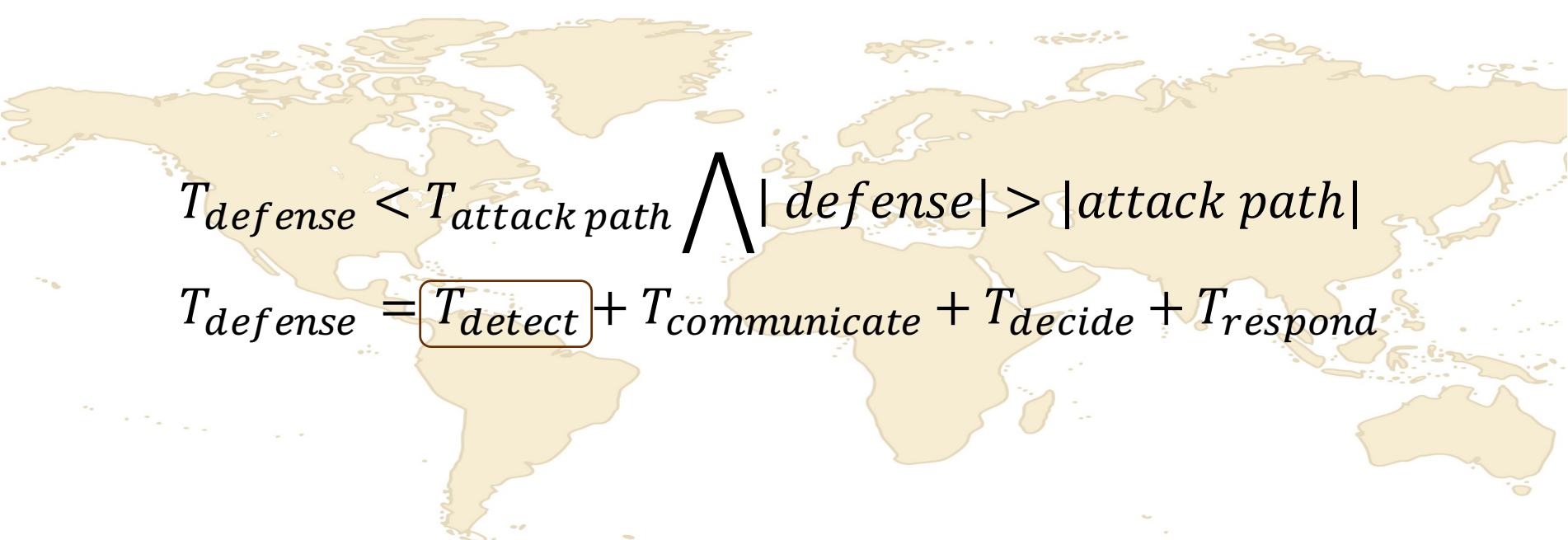
Attacks are through path in
a graph



Attacks are through path in
a graph



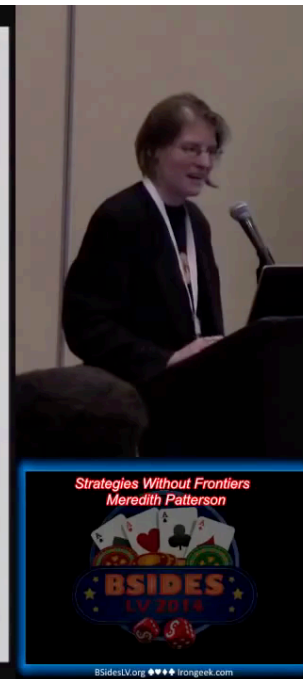
Our Goal as Defenders (for a single attack path)


$$T_{\text{defense}} < T_{\text{attack path}} \bigwedge | \text{defense} | > | \text{attack path} |$$
$$T_{\text{defense}} = \boxed{T_{\text{detect}}} + T_{\text{communicate}} + T_{\text{decide}} + T_{\text{respond}}$$

When an unknown agent acts...

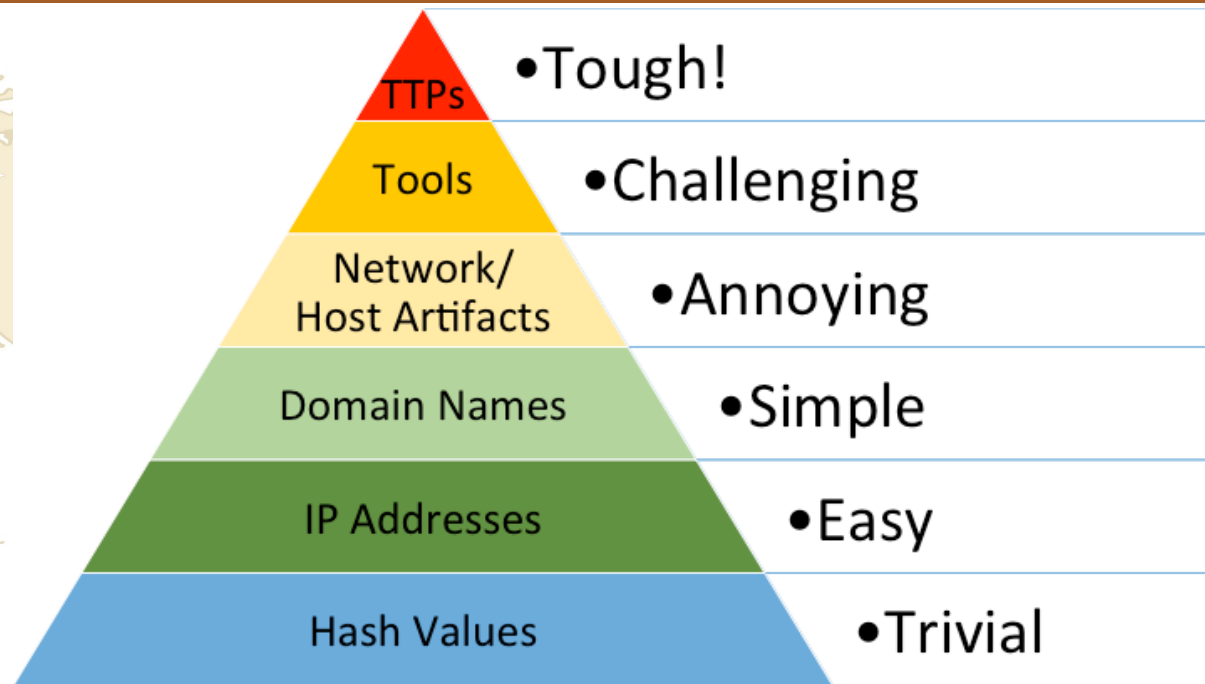
IT IS PITCH BLACK. YOU ARE LIKELY TO
BE EATEN BY A GRUE.

■ When an unknown agent acts, how do you react?



<https://youtu.be/jWxtTsRJOYg>

Not all indicators are equal



<http://detect-respond.blogspot.com/2014/03/use-of-term-intelligence-at-rsa.html>

Infrastructure Indicators

DEF CON 22 - Mahjoub, Reuille, and Toonik - Catching Malware En Masse: DNS and IP Style

IP style

Study relationships between ASNs

- 3100+ malware domains on 1020+ IPs !
- Payload URLs were live on entire IP ranges before any domains were hosted on them
- Seems the IP infrastructure is set up in bulk and in advance

<http://pastebin.com/X83gkPY4>

22 DEF CON

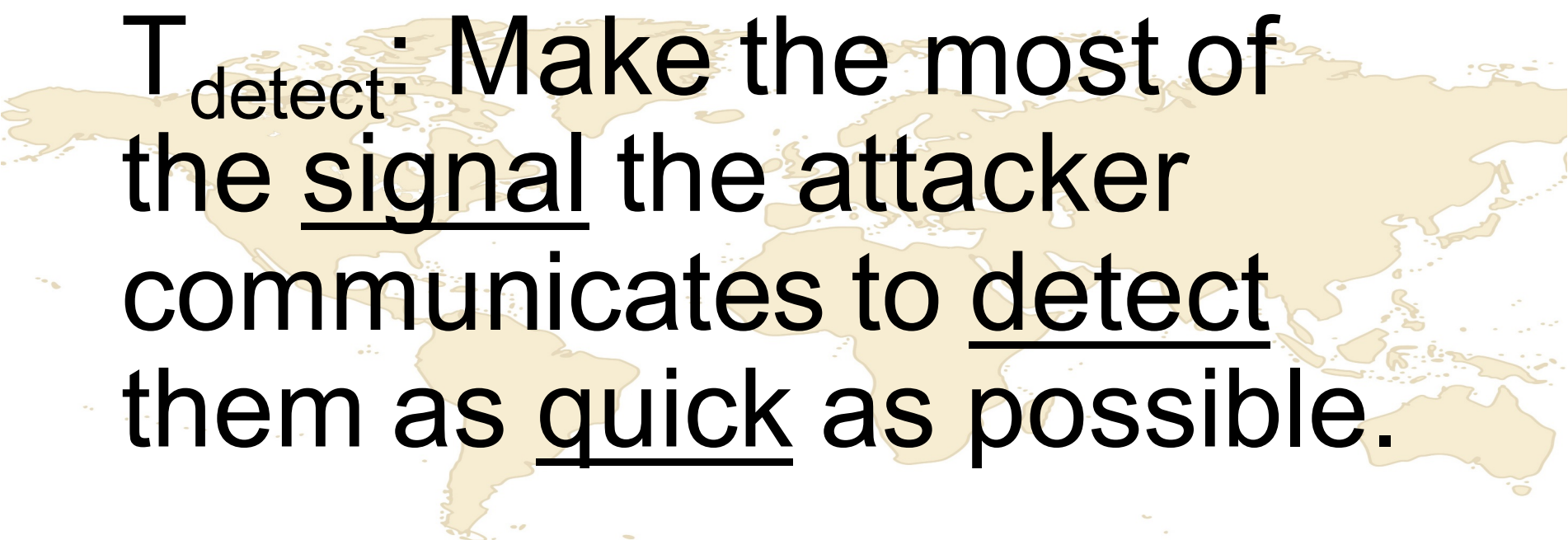
DEF CON

37:30 / 40:15

<https://www.youtube.com/watch?v=KFx4lhxMi-M>



Our Job

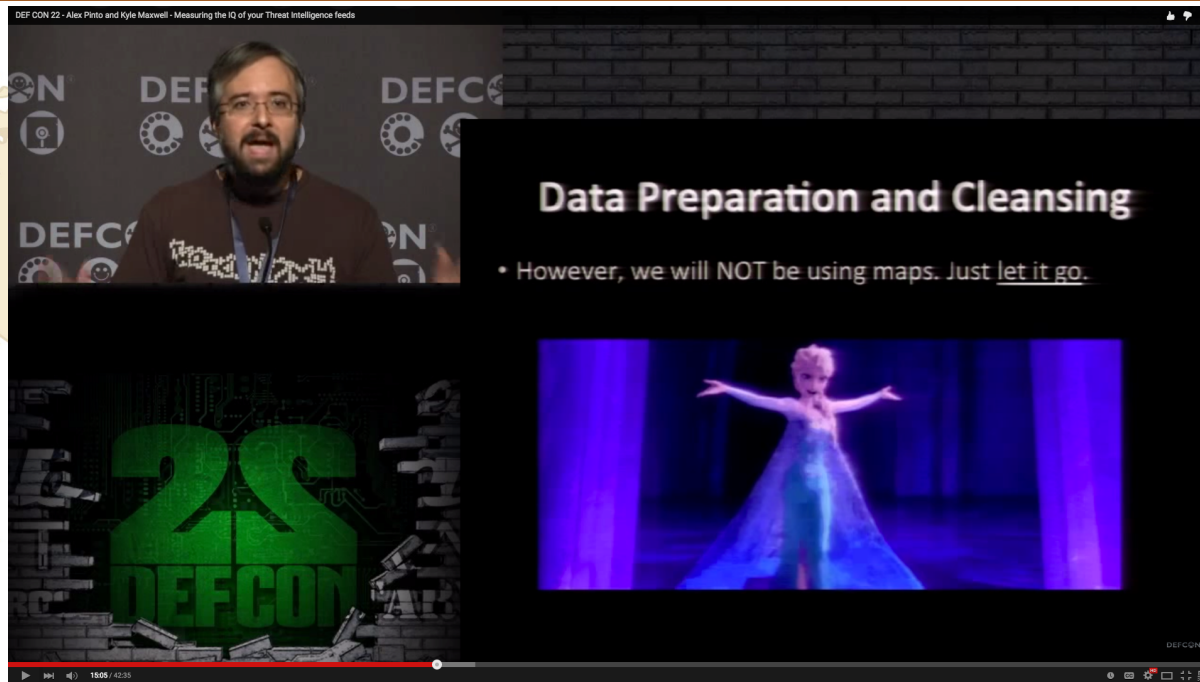


T_{detect} : Make the most of the signal the attacker communicates to detect them as quick as possible.



Why threat intel sucks

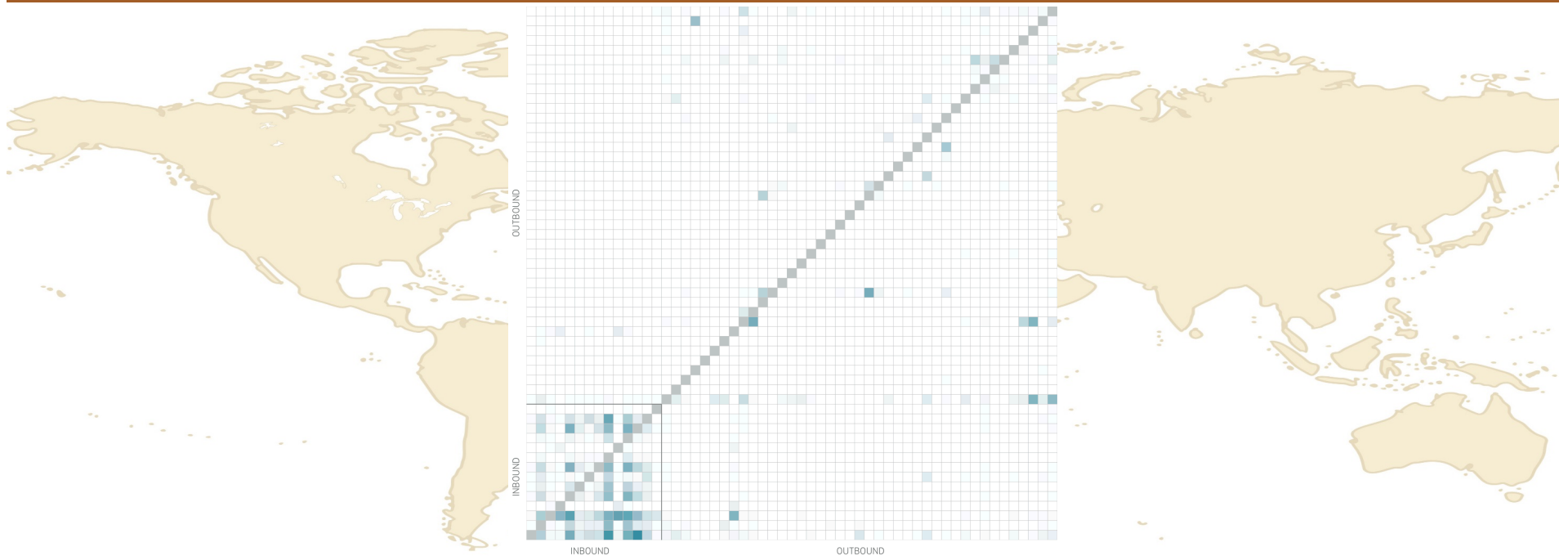
I'm not the first...



<https://github.com/nlsecproject/tiq-test>

<https://github.com/mlsecproject/combine>

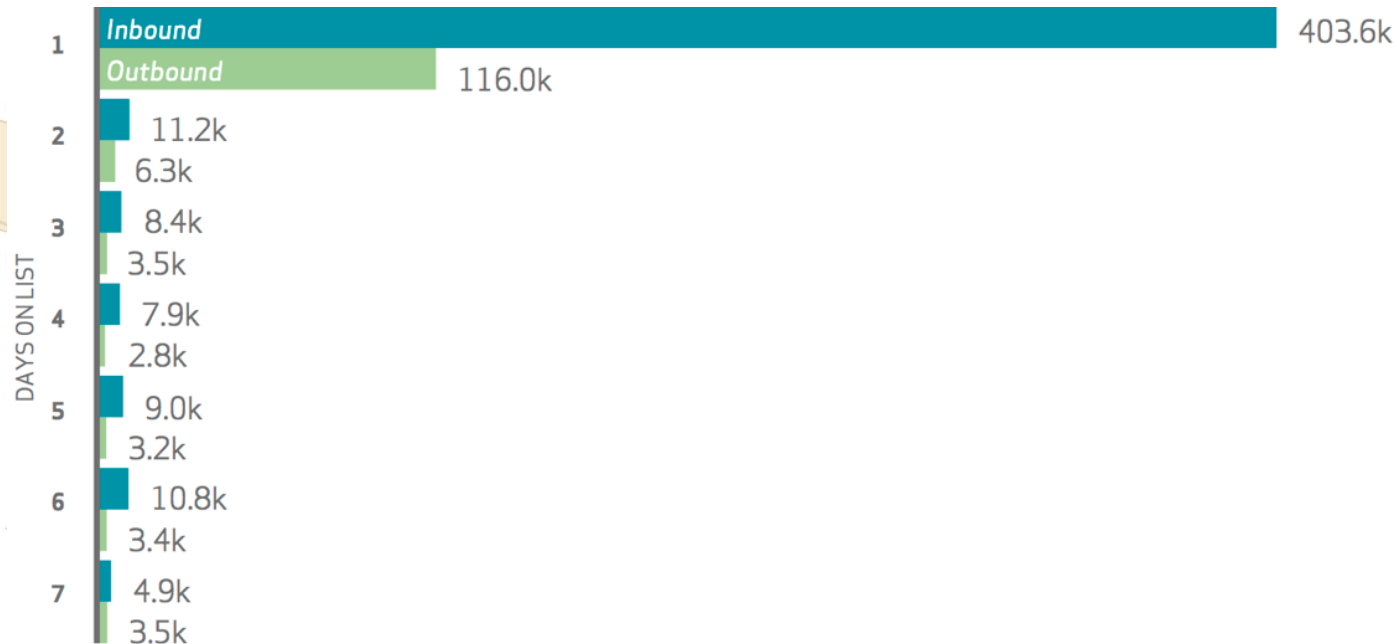
Intel is unique: 3% overlap



2015 Verizon Data Breach Investigations Report

<https://securityblog.verizonenterprise.com/?p=6848>

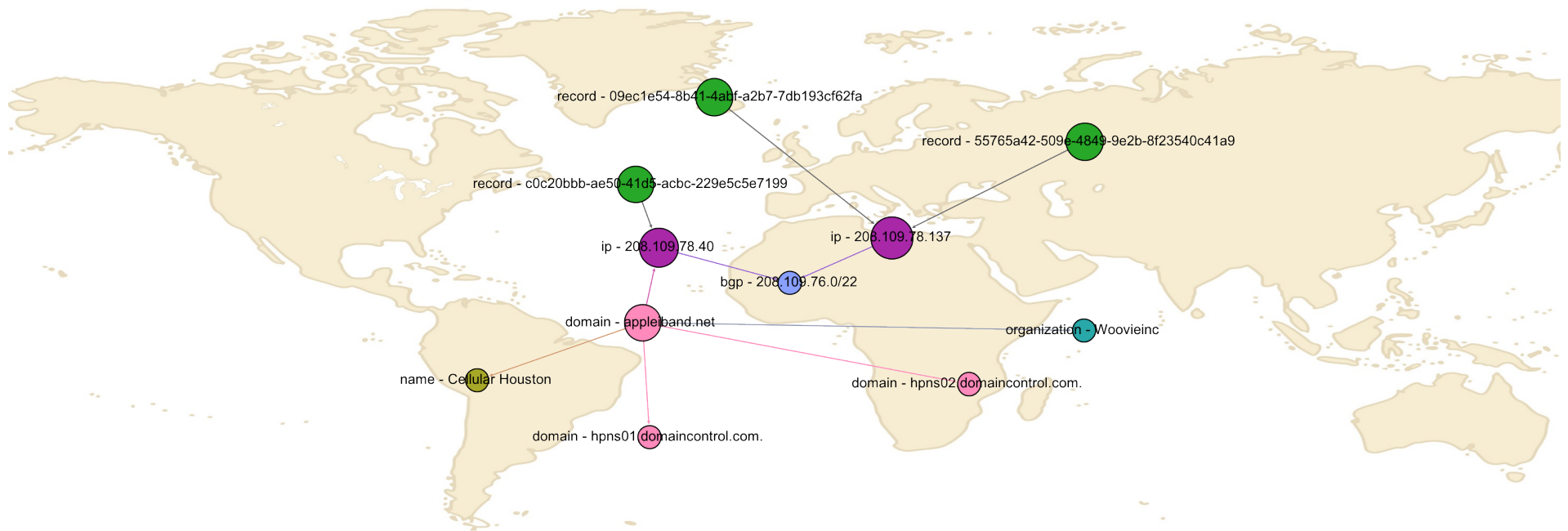
Indicators Burn Fast



2015 Verizon Data Breach Investigations Report

<https://securityblog.verizonenterprise.com/?p=6848>

But what if we knew more...



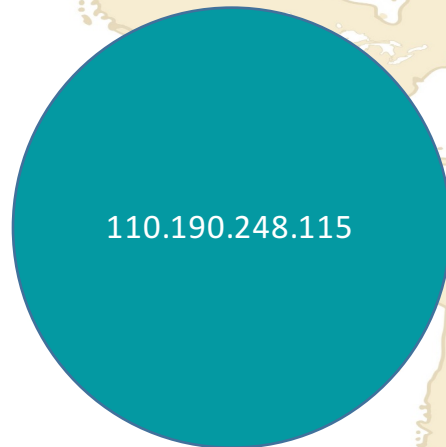


Graph Theory 101 (abbreviated)

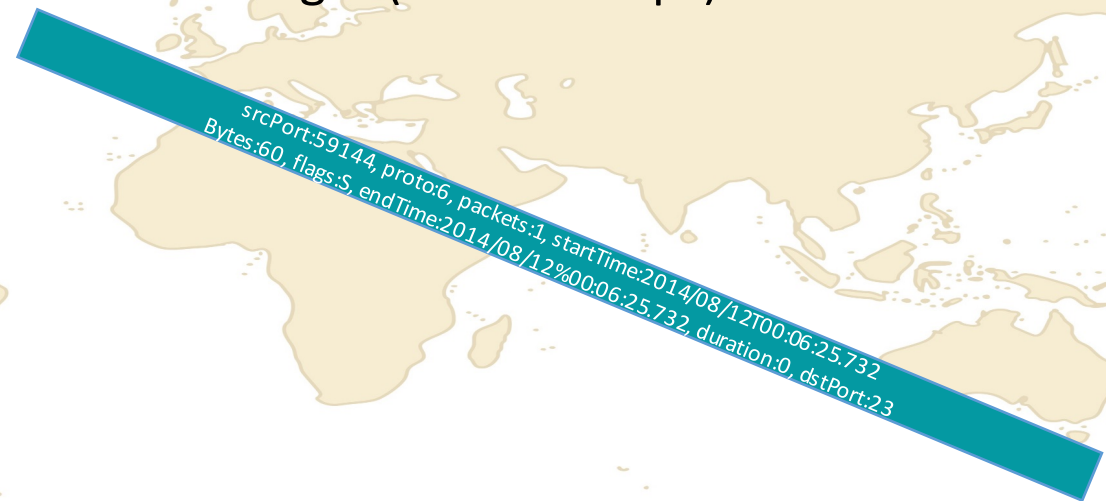
What's a graph?

A collection of:

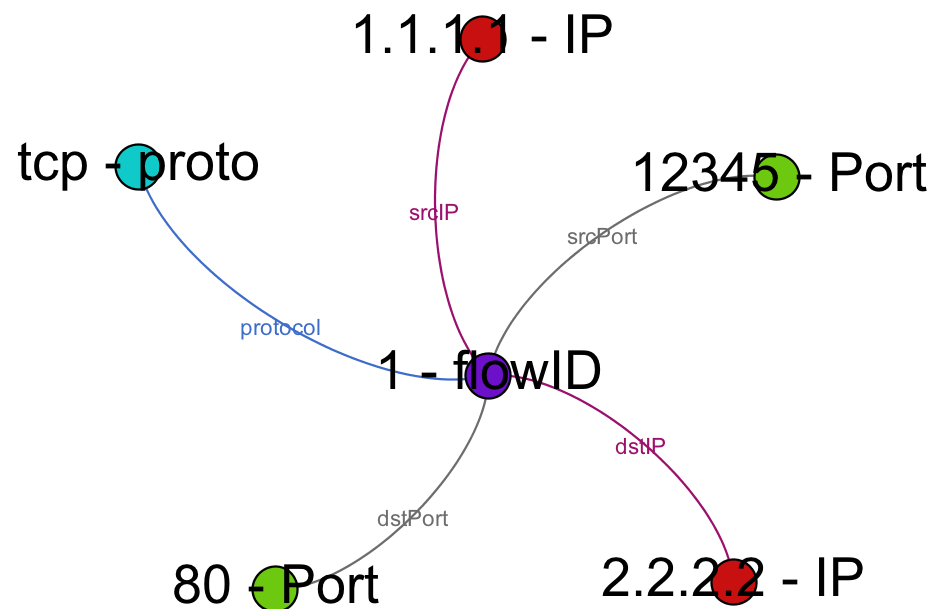
- Nodes (Vertices)



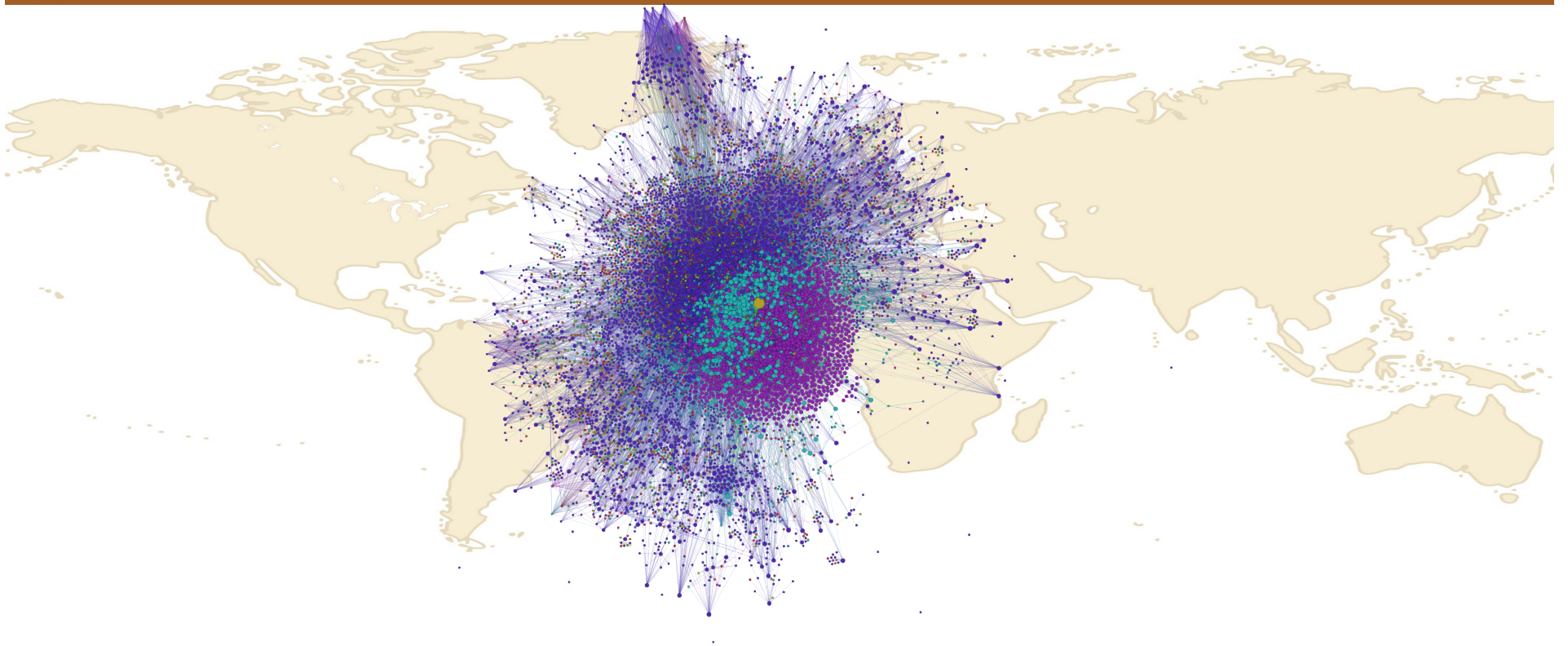
- Edges (Relationships)

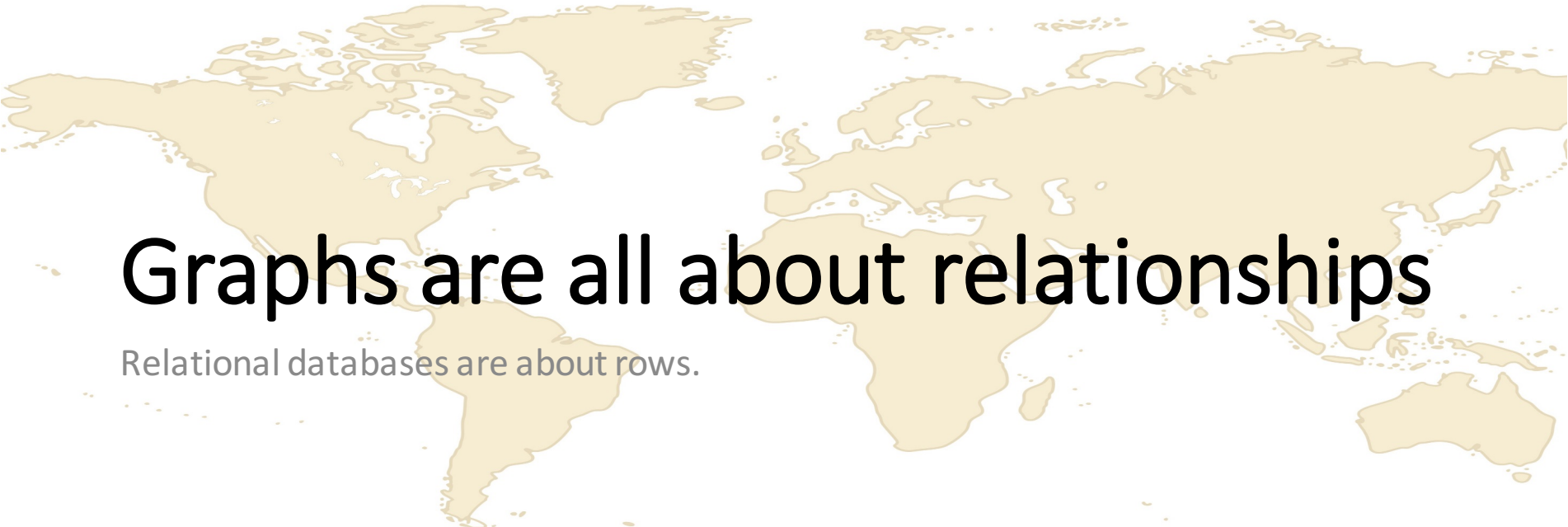


What's a graph?



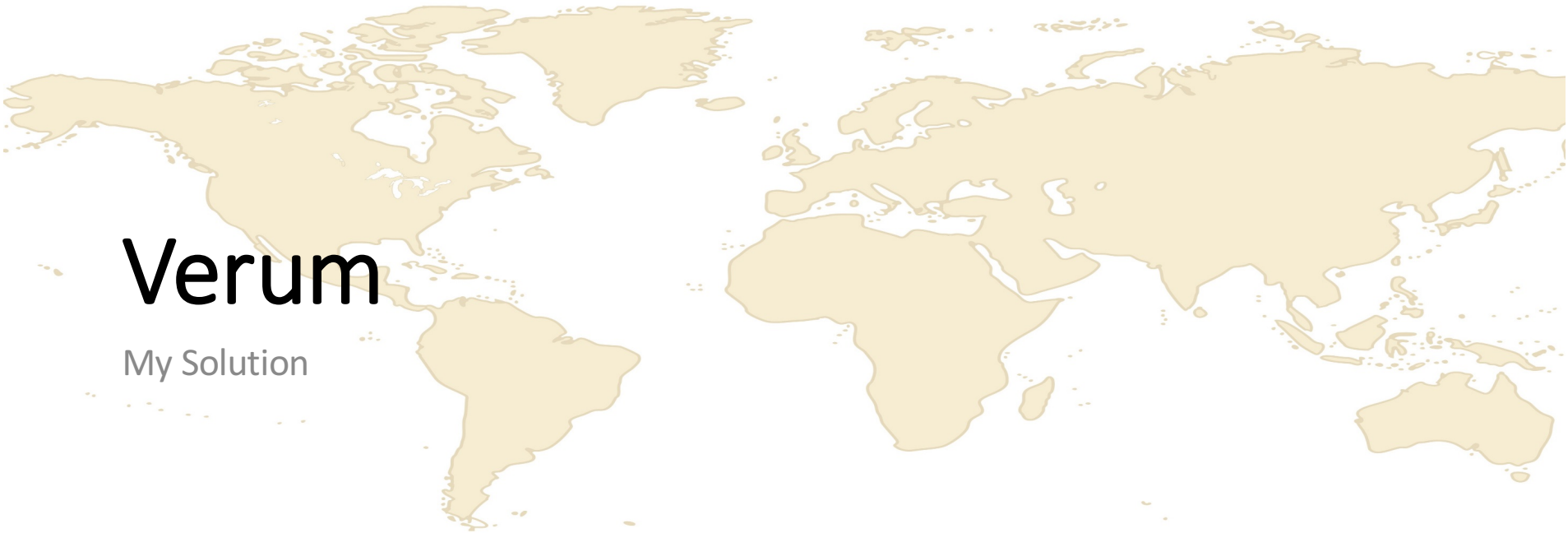
Not about the looks...





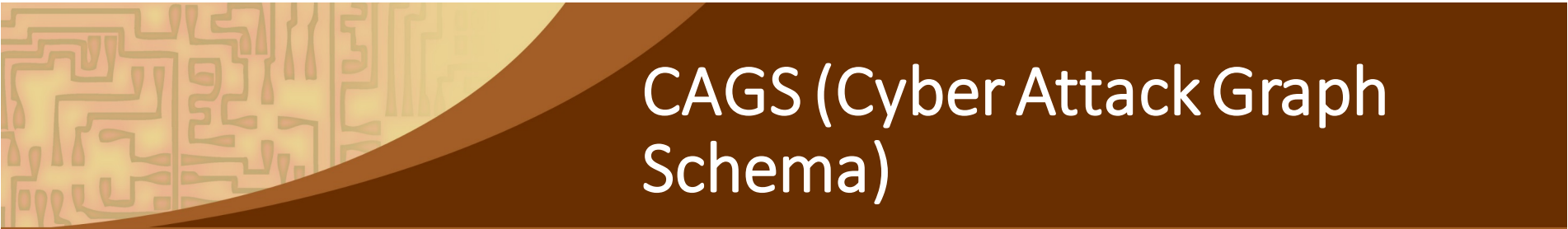
Graphs are all about relationships

Relational databases are about rows.

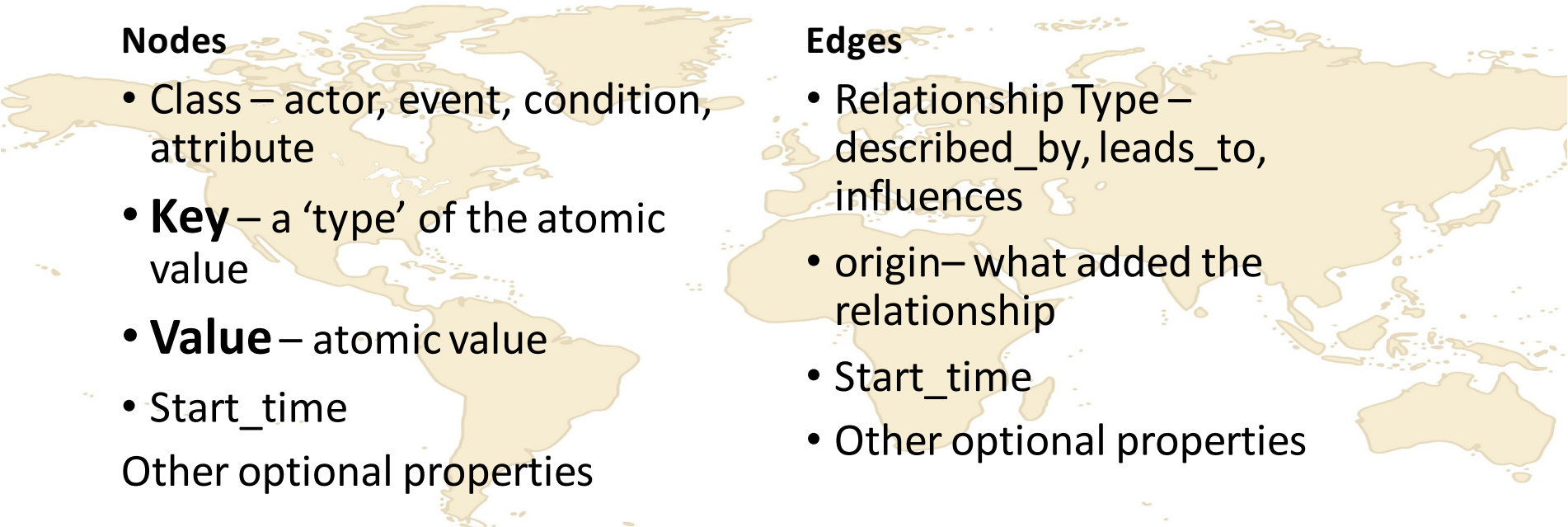


Verum

My Solution



CAGS (Cyber Attack Graph Schema)



Nodes

- Class – actor, event, condition, attribute
- **Key** – a ‘type’ of the atomic value
- **Value** – atomic value
- Start_time

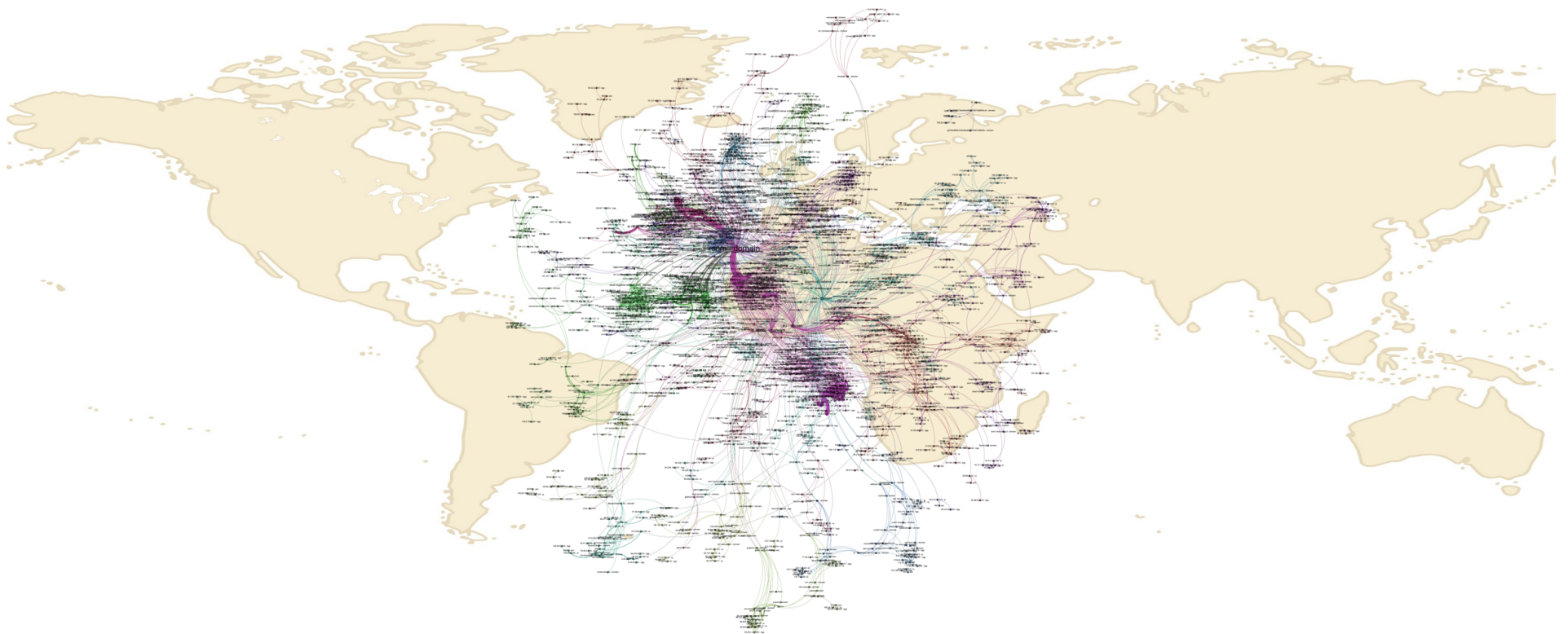
Other optional properties

Edges

- Relationship Type – described_by, leads_to, influences
- origin– what added the relationship
- Start_time
- Other optional properties

<http://blog.infosecanalytics.com/2014/11/cyber-attack-graph-schema-cags-20.html>

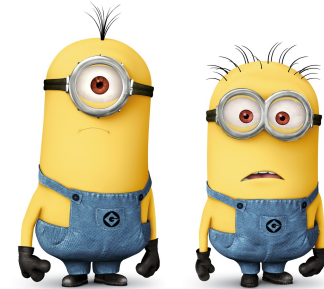
CAGS



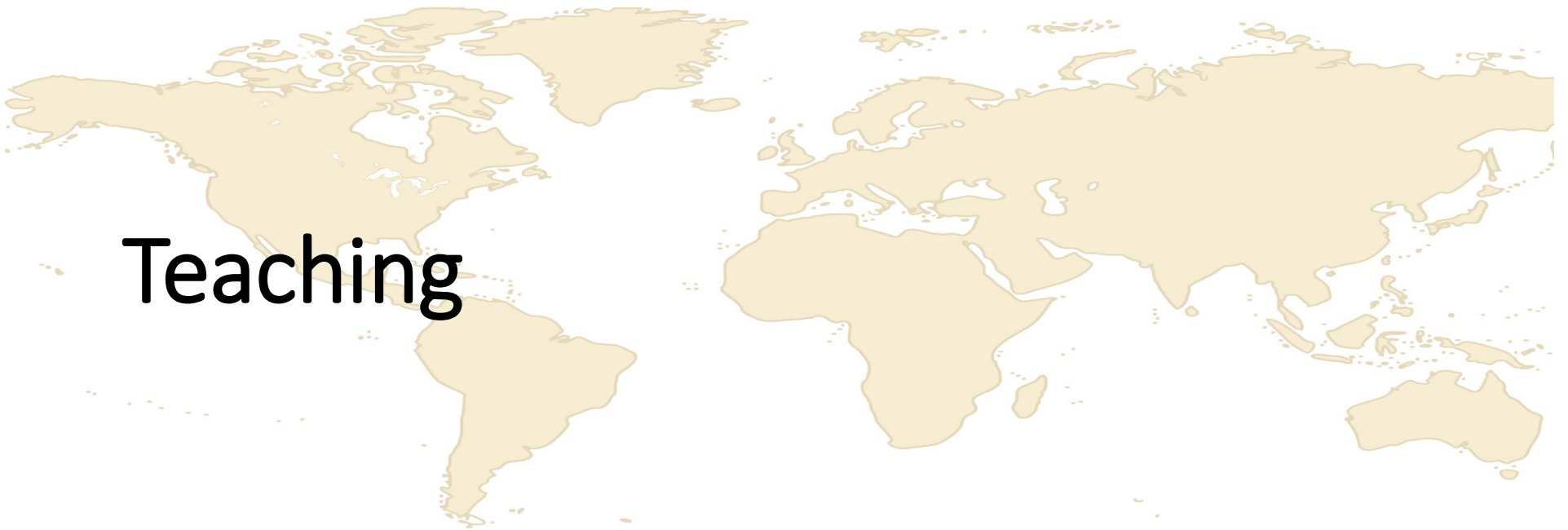
Implementation

- A simple arbiter which loads plugins
 - Enrichment
 - Interface (storage)
 - Scoring
- Minions
- Written in Python using YAPSY for plugin management

Backend Storage:
Networkx
Neo4j
TitanDB



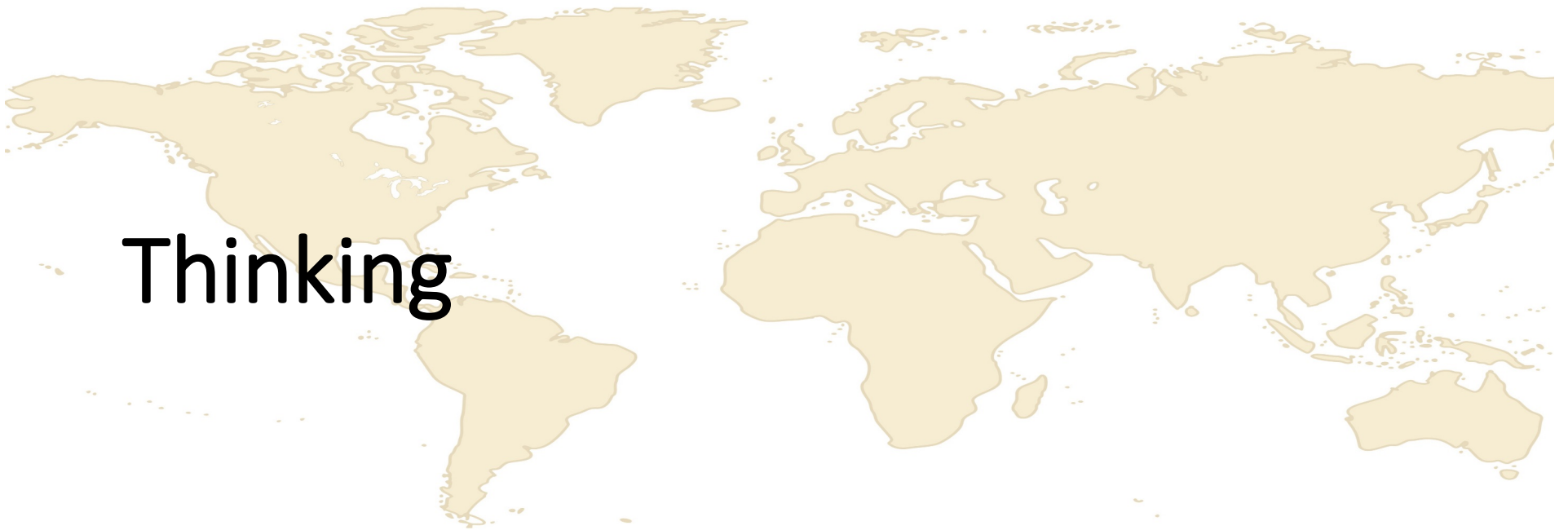
Storage backend becomes primary bottleneck though scoring can be as well.



Teaching



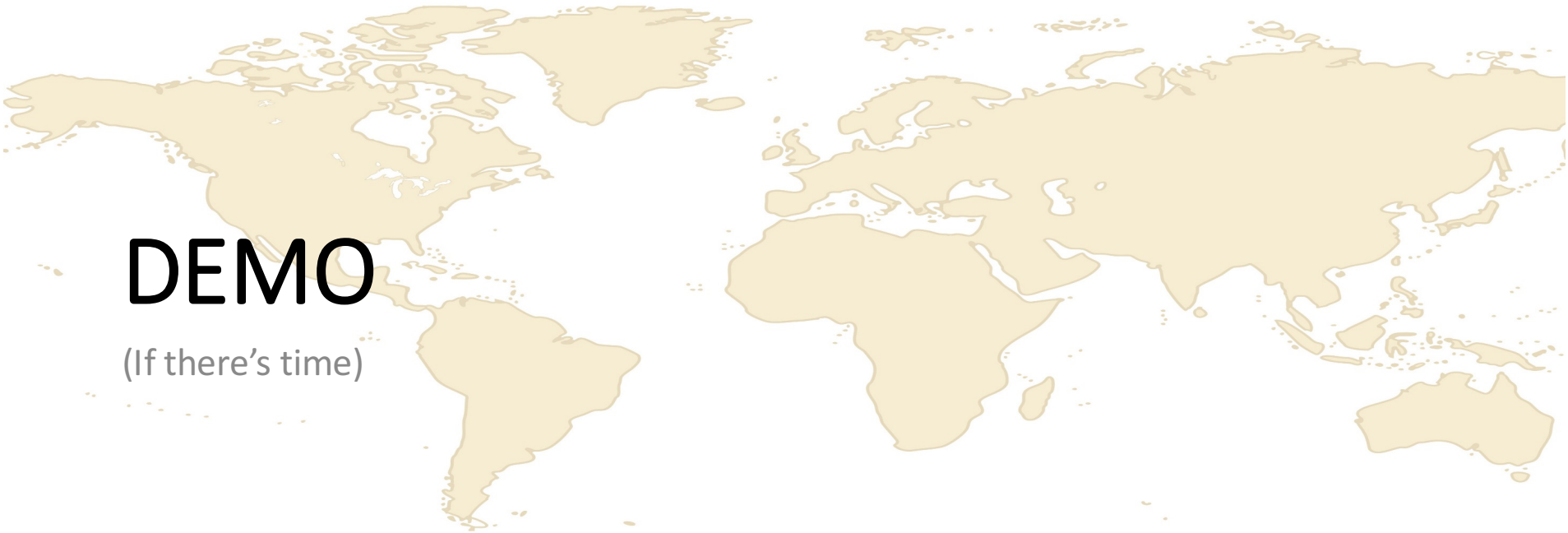
Remembering



Thinking



Communicating

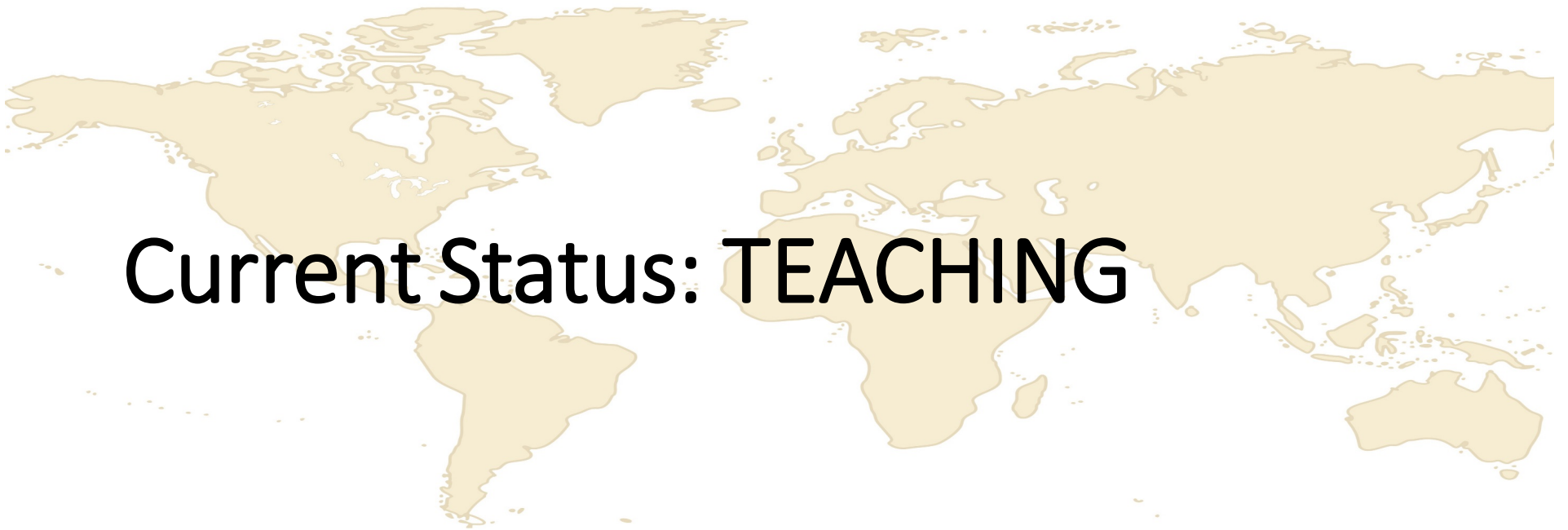


DEMO

(If there's time)



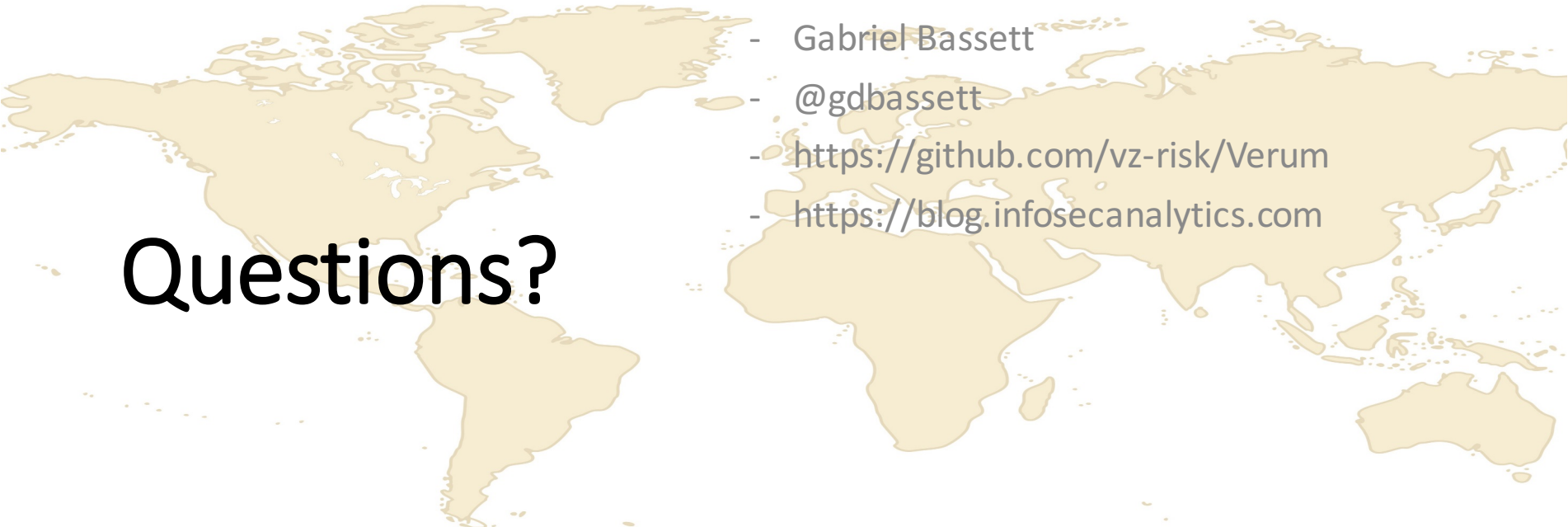
Not your normal classifier!



Current Status: TEACHING



Conclusion



Questions?

- Gabriel Bassett
- @gdbassett
- <https://github.com/vz-risk/Verum>
- <https://blog.infosecanalytics.com>