



AUGIFY

The Future of Understanding



Cloud-based Data Science & Interactive Narratives

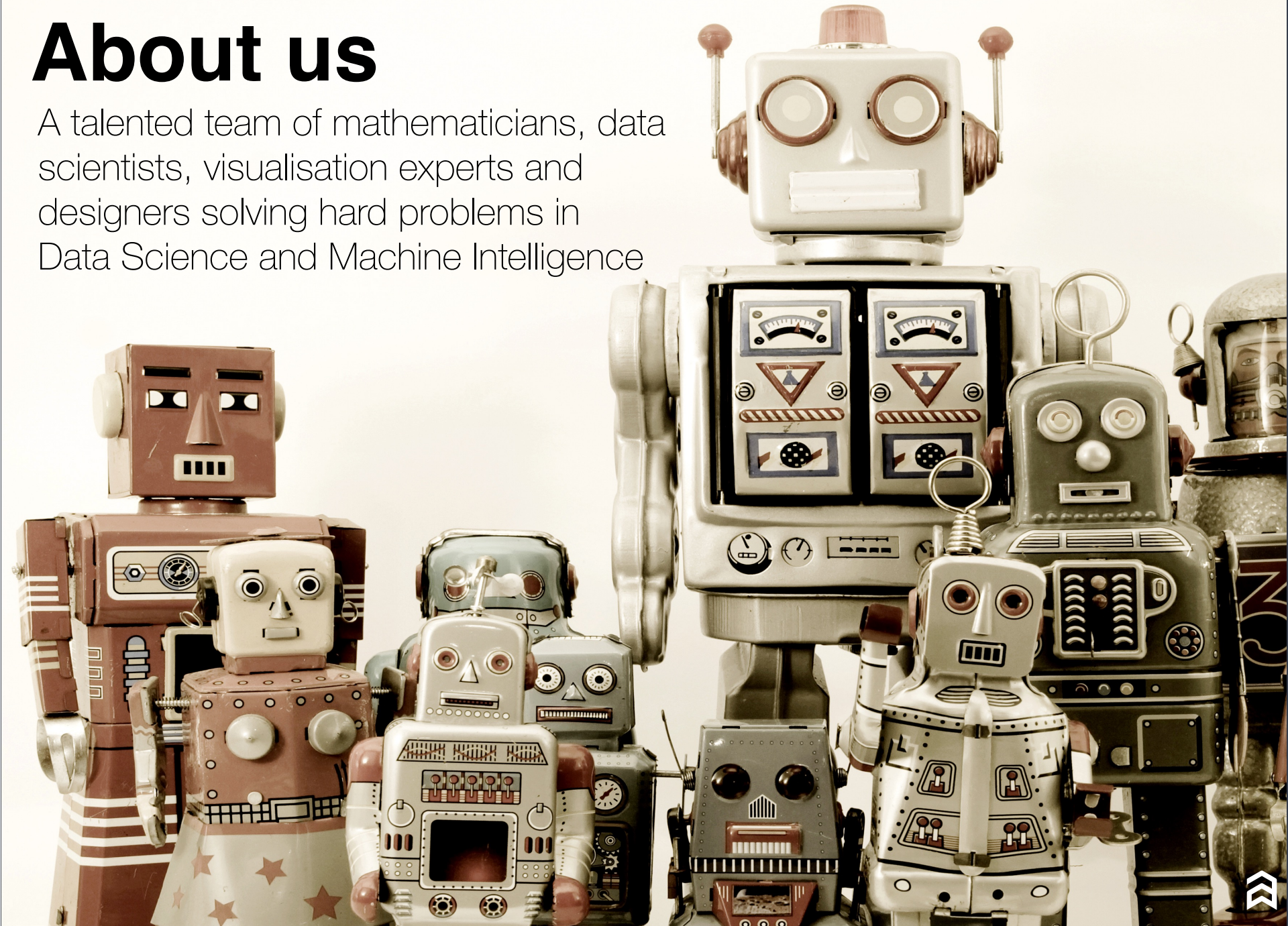
Jay Solomon

Founder and Chief Science Officer

August 2015

About us

A talented team of mathematicians, data scientists, visualisation experts and designers solving hard problems in Data Science and Machine Intelligence



Clients

Complex Data Science at Massive Scale







Social Media and Mainstream Media



Large-scale News and Media Monitoring



Patent Analysis and Network Detection



Intellectual Property Theft Detection



Political Network and Influencer Analysis



Hyper Immersive Visual Environments



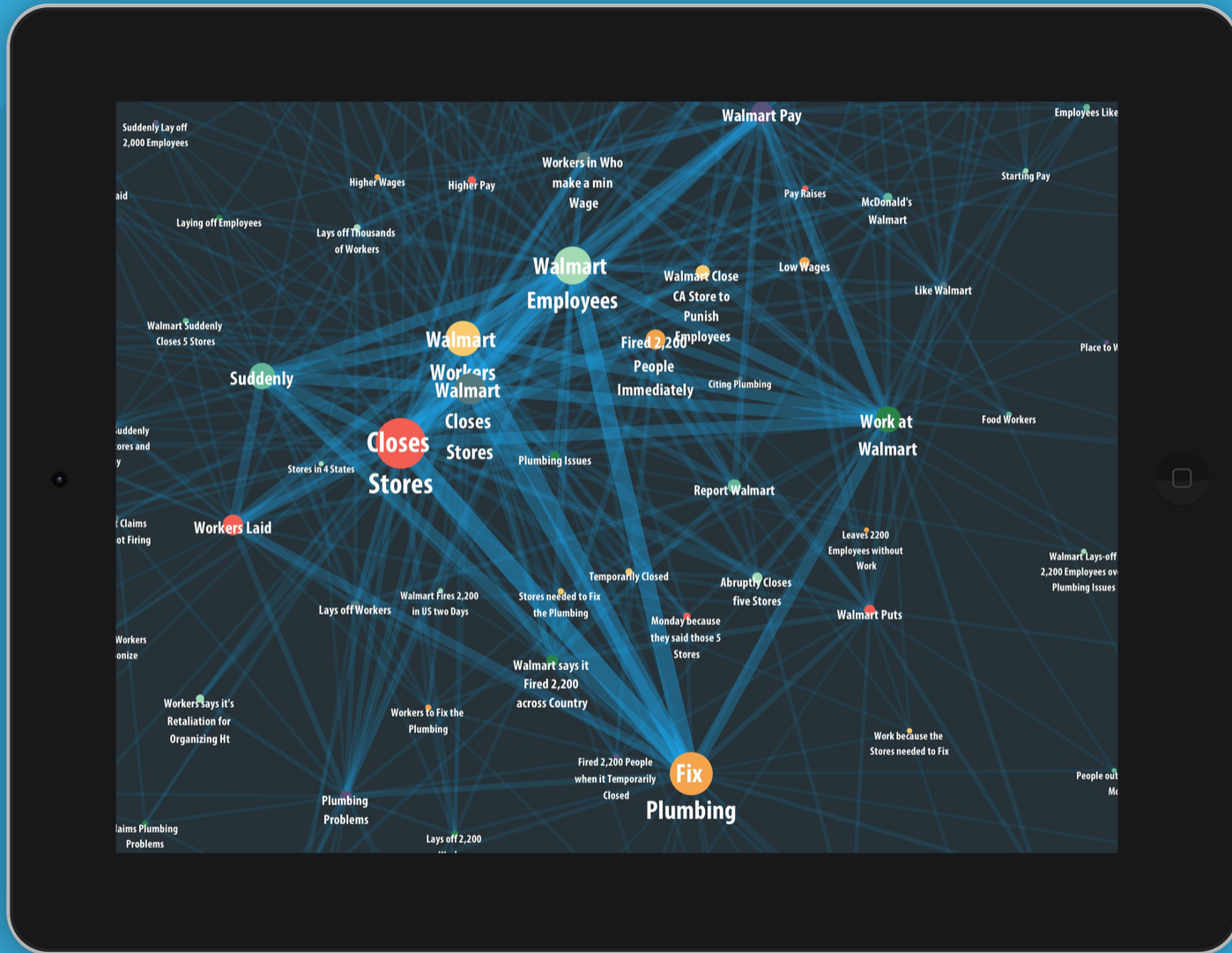
Cyber Security Threat Vector Analysis



Crisis Management, Political Network Analysis, Social Listening, Mainstream News, Competitor Intel



MacBook Pro

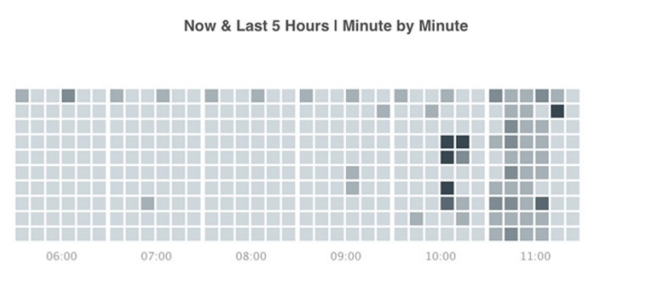
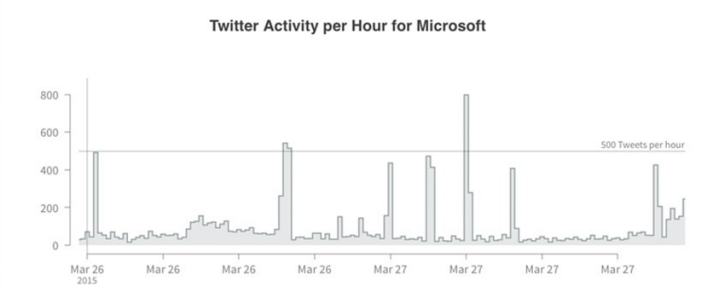


Interactive Mobile Big Data Applications: Conversation Topic Networks for Crisis Detection

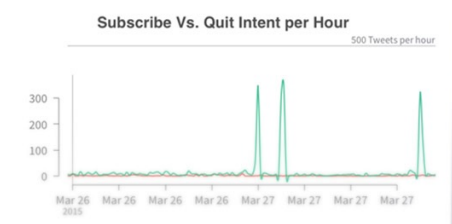
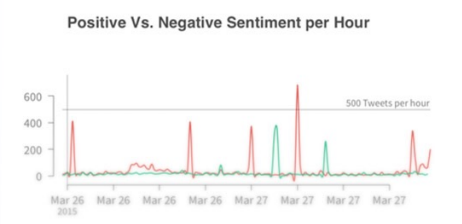


- 1 #MICROSOFT Insights or Networks
- 2 #TECH Insights or Networks
- 3 #AZURE Insights or Networks
- 4 #APPS Insights or Networks
- 5 #GAMES Insights or Networks
- 6 #SONY Insights or Networks
- 7 #LGMAGNA Insights or Networks
- 8 #LUMIA Insights or Networks
- 9 #APP Insights or Networks
- 10 #LAPTOP Insights or Networks
- 11 #SKYPE Insights or Networks
- 12 #HOWTO Insights or Networks
- 13 #YOSEMITE Insights or Networks
- 14 #FOLLOW Insights or Networks
- 15 #LOVE Insights or Networks
- 16 #1D Insights or Networks
- 17 #MIND Insights or Networks
- 18 #ZAYN Insights or Networks
- 19 #ZAYNMALIK Insights or Networks
- 20 #SOCIAL Insights or Networks
- 21 #BBK Insights or Networks
- 22 #BBK Insights or Networks

24 Hours Tweets tracked	Total Tweets In Capgemini Stream	Yesterday Microsoft	Today Microsoft	Change In 24 Hours	Velocity Twitter Activity
186.9K	1.4M	7.5K	6.4K	23%	HIGH



Positive Sentiment Tweets tracked	Negative Sentiment Tweets tracked	Buy Intentions Tweets about buying	Sell Intent Tweets about Selling	Subscribe Intent Tweets about subscribing	Quit Intent Tweets about leaving
23.8%	47.8%	40.1%	21.2%	19.5%	1.5%



Rumours Rumour and untruth	Social Ties Shared and Connected	Anxiety An anxious tone	Personal Connection From a personal perspective	Desire Tweets signaling desire	Growth Improvment and growth
17%	30.2%	0.1%	5.5%	1.6%	%



TOPIC IDENTIFICATION AND NETWORKS FOR WALMART

Detecting topics discussed in social conversations and their interconnections

Immersive and Interactive Big Screen Dashboard:

Unified Dashboard for Social, TV, News, Web Data

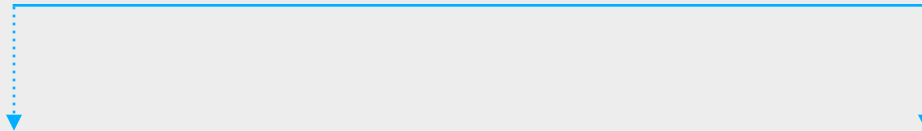


Products

Science / Narratives / Streams



2 x Cloud Platforms



 AUGIFY

Science

Data Science & Deep Learning

- Data Harvesting
- Data Science
- Deep Learning
- Prediction
- Search Analytics

 AUGIFY

Narrative

Interactive & Immersive Storytelling

- Advanced Visualisation
- Automated Insights
- Immersive Experiences
- Micro to Massive Scale
- Generative Narratives

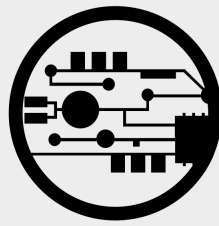


Data Streams

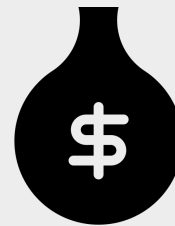
Real-time indexes of human and machine generated data from social, news, chat, forum, patent, and 100s of data sources intelligently augmented with meaning.



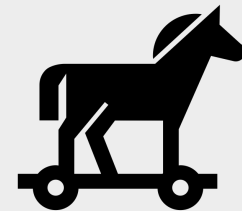
WorldIndex



TechIndex



MarketIndex



CyberIndex

New index possibilities or custom client indexes

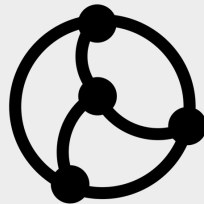


Science

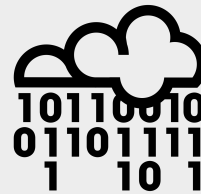
Harvest Data in real-time from structured and unstructured sources inside and outside of the firewall



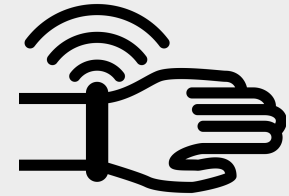
**Enterprise
Data**



**Public
Data**



**Open
Data**



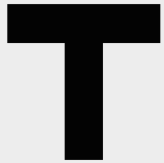
**Event
Data**

1000s of connectors to data sources

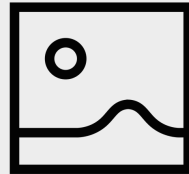
 AUGIFY

Science

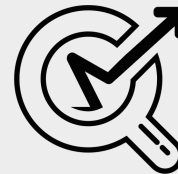
Add Meaning to data in real-time using a collection of data science modules and intelligent algorithms



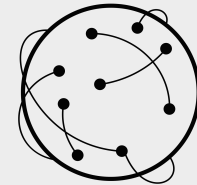
**Text
Analytics**



**Image
Analysis**



**Search
Analytics**



**Network
Science**

100s of algorithms and modules

Algorithms and Data Science Modules

Intentions	Things	Networks
Political Emotions	Topics	Events
Financial Emotions	Rumour	Summaries
Sentiment	Truth	Stylometry
Financial Keyphrases	Speculation	Influence
Lifescience Keyphrases	Credibility	Demographics

Search
Search Analytics
Prediction
Graph Analytics
Deep Learning



Narrative

400+ *Visualisations*

Tell Stories with data and understand complex fast changing scenarios. Create usable experiences on screens that make intelligence actionable.



Flows



Time



Networks



Proportions



Hierarchies



Maps

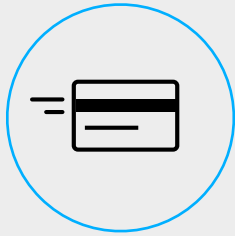


Statistics



Images

A flexible subscription service



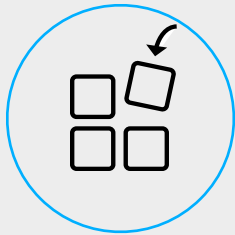
Pay for what you use

Subscription service for Augify Science and Augify Narrative Cloud Platforms.



In your browser

No software installation required.
Multiple user accounts.



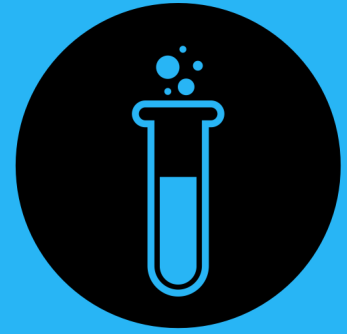
Choose components

Choose the data stream, data science and narrative components you need.



Powered by Augify

White label applications.



What are we working on?

Viral Event Prediction

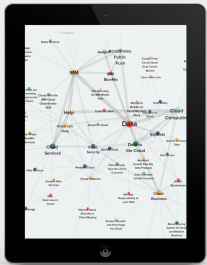
Temporal Event Causality

Text Understanding in Many Languages using
Deep Learning

Audio Transcript Analysis using Deep Learning

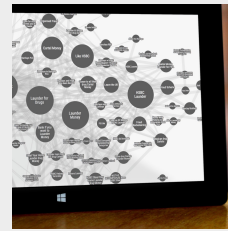
Massive Graph Visualisation

Current Client Challenges in Data Science



Massive Graph Visualisation

Visualise massive graphs and provide new ways to filter, detect communities, and compare time slices



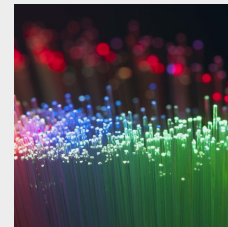
Virality & Meme Identification

Identify and understand the dynamics of emerging topics and memes connected to a topic, entity, artist or music genre



Emotion Classification

Classify emotions using Deep Learning and gain a deep understanding of text



Event Causality

Predict what people will talk about based on temporal events



Finding the right balance

From research to commercially viable

**Balance innovation,
computational cost, and
commercial viability**



1

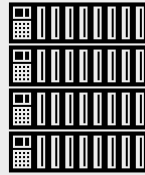
Data Science



Leverage existing research and tools
Vs. create new paradigms

2

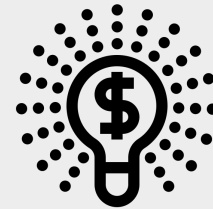
Computational Cost



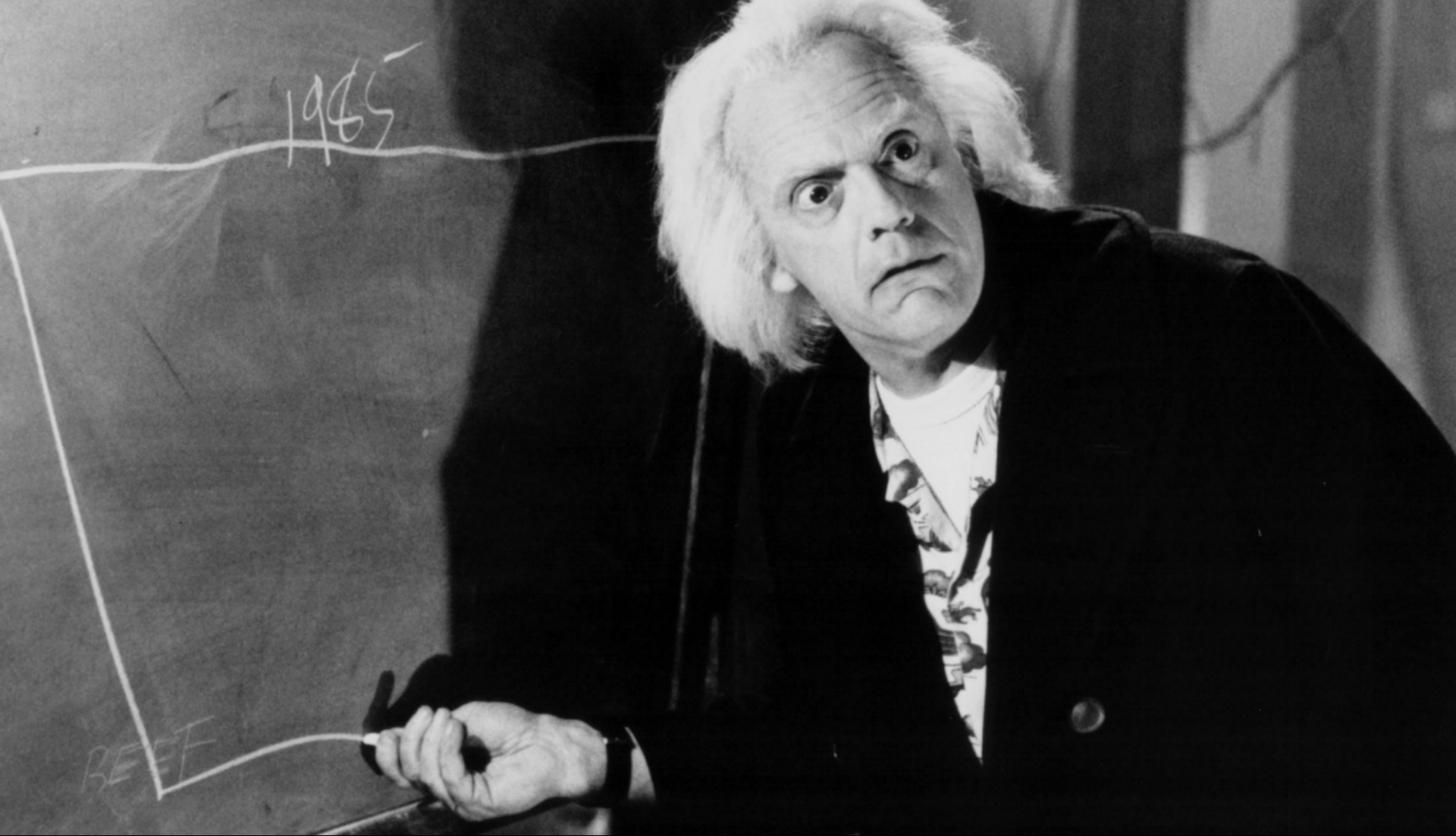
Cloud Infrastructure
Training Deep Learning Models
Massive Scale

3

Commercial Viability



Mass-market low entry price Vs. Specialist high value



Challenge

Large Scale Graph Visualisation and Interaction

Big Graph Visualisation Challenge

Find efficient and innovate ways to make large amounts of data legible and scale visual design for Big Graphs with 20,000+ nodes

How are we using Big Graphs? The data

- Social Media and Mainstream Media Analysis
- Large-scale News and Media Networks
- Patent Analysis and Network Detection
- Intellectual Property Theft
- Political Network and Influencer Analysis
- Hyper Immersive Visual Environments - Massive screens
- Cyber Security Threat Vector Analysis

Scaling Visual Design with nodes, edges, colors, sizes, and filtering



1

Nodes

Automatically balance
Attraction/Repulsion
forces

2

Edges

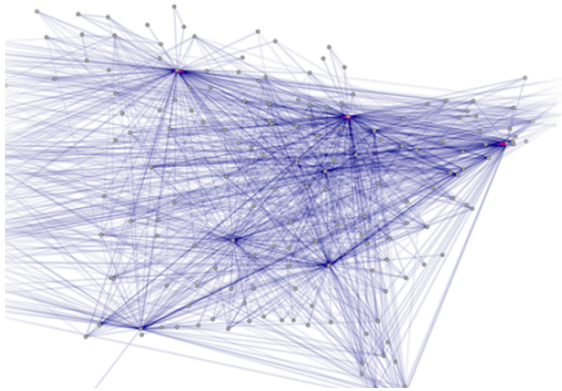
Arcs
Bundled Edges

3

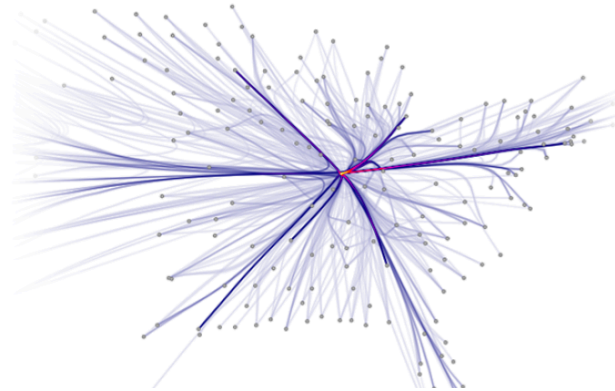
Size/Color

Algorithms and
Graph Metrics for
Community Detection
and Centralities

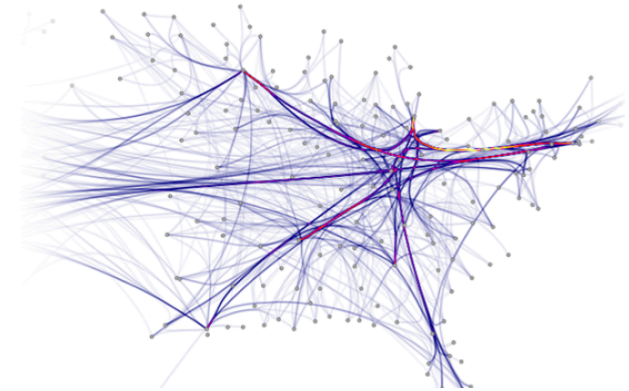
Challenge find efficient ways to perform edge bundling on large graphs. Client vs. Cloud/GPU techniques should be explored.



No Bundling

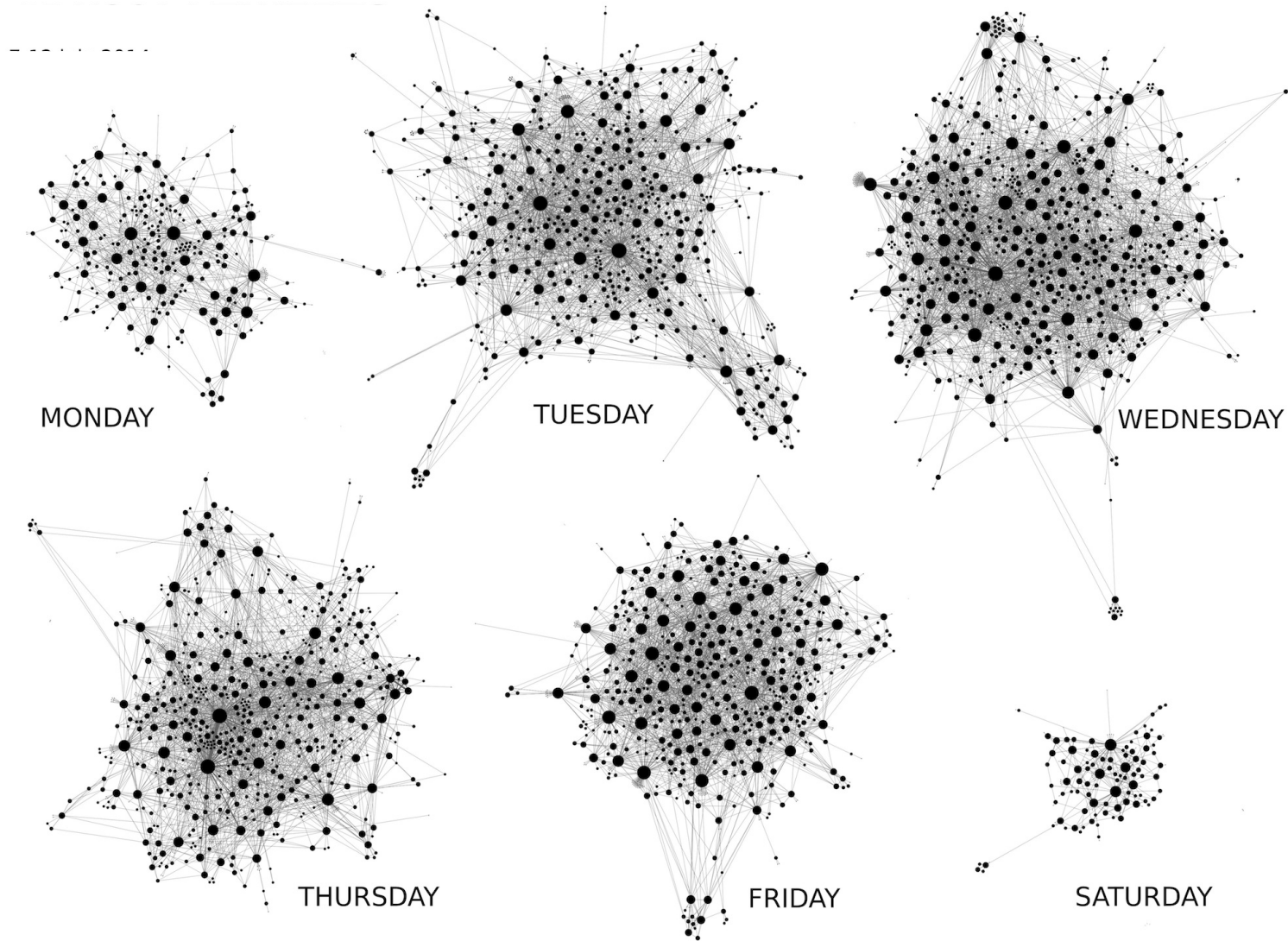


All Edges Bundling

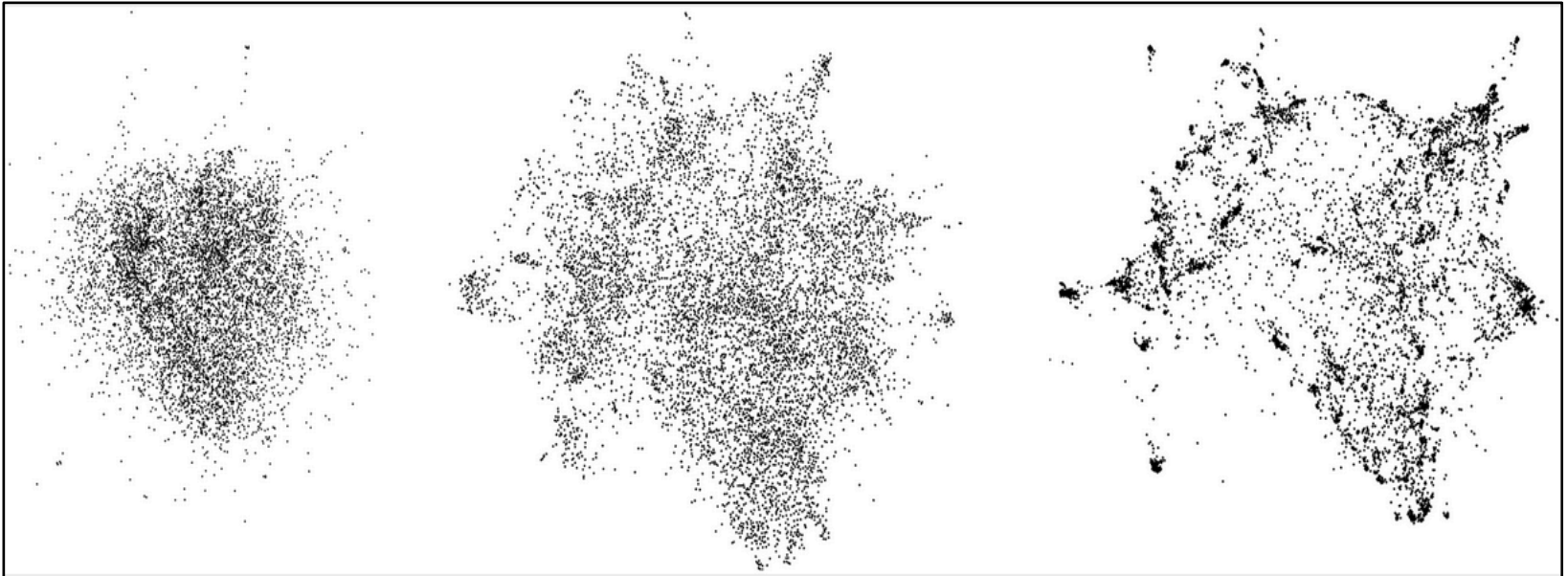


Smart Bundling

Challenge visualise and compare multiple graphs for each day and provide live filters for edges, communities, outliers, metrics all updated in 1 view



Challenge automatically balance attraction and repulsion forces for each graph based on data for each query / network



Repulsion	$1 / \text{distance}$	$1 / \text{distance}$	$1 / \text{distance}$
Attraction	distance^2	distance	$\log(\text{distance})$

1

Betweenness

Bridges across
communities

2

Closeness

Well connected

3

Influence

PageRank

Challenge calculate metrics in real-time.

Challenge detect communities and outliers.

Computing additional information

- Node community (Infomap, Louvain)
- Edge community (Link Communities)



[Communities in Networks, Porter et al., Link Communities, Ahn et al]

Challenge Filter the network in real-time

Visibility ✕

X selected node

Select which hastags and relations to display.

Prune Hide

Hashtag Degree: 0

Relations min weight: 3

link dist factor (pagerank) [0-100] 1

link dist factor (weight) [0-10] 1

node size (Degree factor) [0-2] 0.7

node size (PageRank factor) [0-2] 0.3

Relations

Relations (bundled)

Nodes

Nodes text

Arcs

Time

Degree

Relations

Link Distance

PageRank

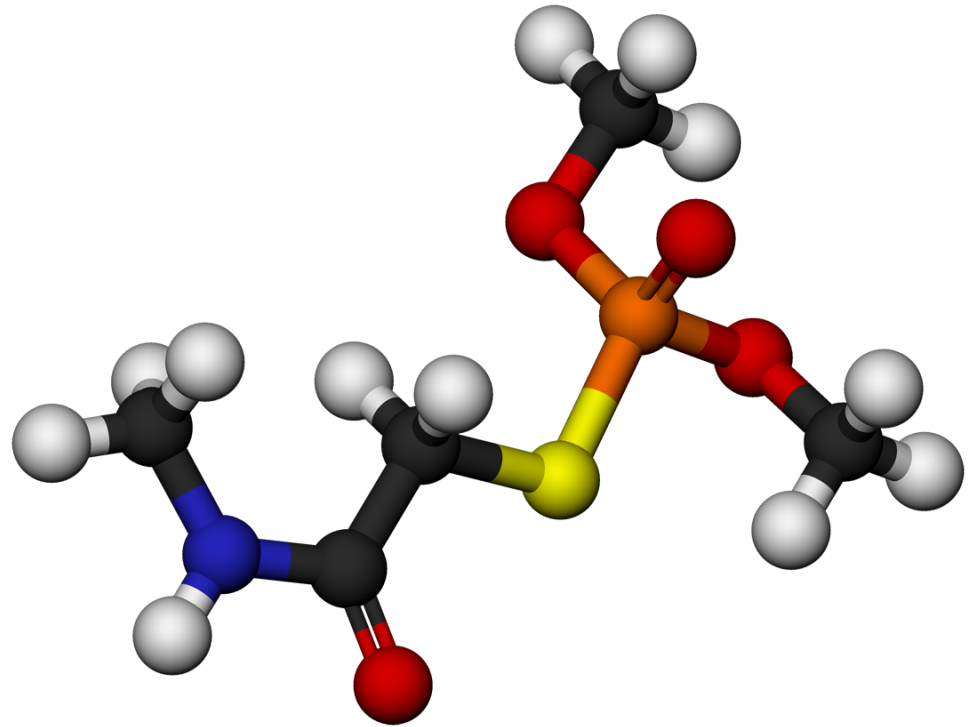
Natural Language Signals

**(Emotions, intentions,
sentiment, events, things)**

Bonus develop interesting ways to visualise networks in VR and 3D



VR



3D

Challenge Summary

Balance Attraction and Repulsion fully automatically

High Performance Edge Bundling from bundling all edges to smart bundling

Colour the Graph

Filter the Graph

Calculate Graph Metrics

Streaming geometry over Wifi/Ethernet

VR and 3D

Tools

D3.js

Cytoscape.js

Webcola



AUGIFY

The Future of Understanding