

HOW NLP SAVES THE WORLD!

From Stopping Terrorist Attacks to Spotting Fake News

Chris Brown & Declan Trezise



Unstructured Data Workflow

COLLECT

80% of data is
UNSTRUCTURED
V-2-T

EXTRACT

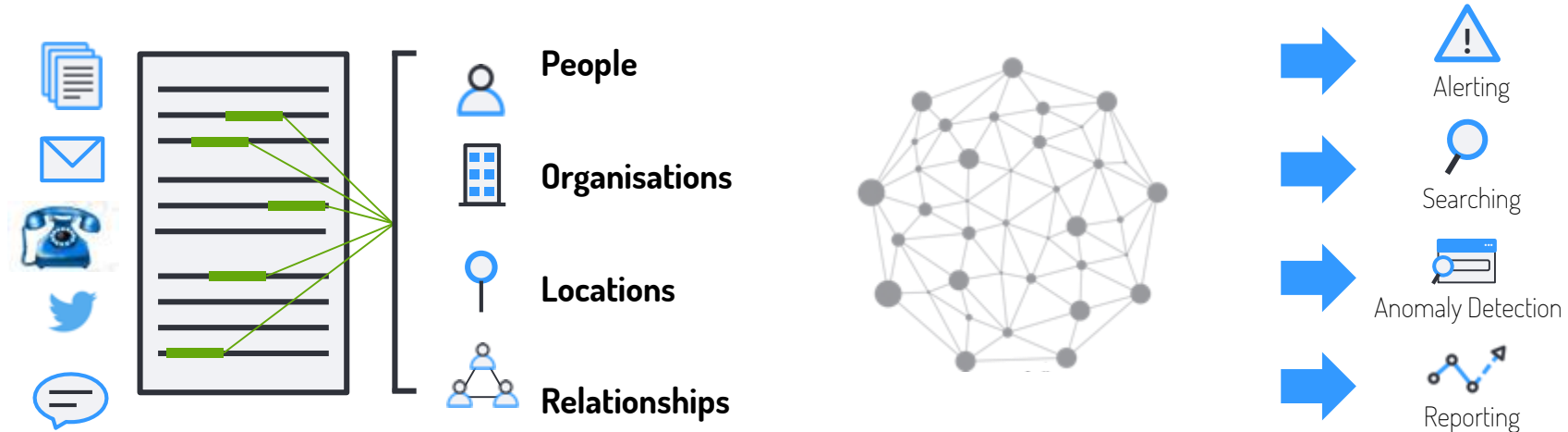
- 1) Natural Language Processing Extracts Facts
- 2) Scored for confidence & relevance

COMBINE

Join Processed and
Structured Data into
Knowledge Graph

ANALYSE

Mine Graph
For Patterns
& Changes



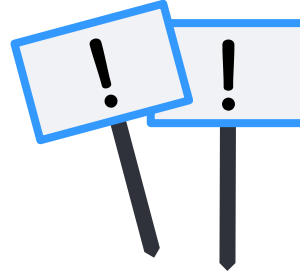
Examples of using NLP



Stops Terrorist Plots



Spots Fake News



Predicts Civil Unrest



Protects Our Borders

Who uses NLP

- Central Intelligence Agency (CIA)
- Defense Intelligence Agency (DIA)
- Department of Defense (DOD)
- Federal Bureau of Investigation (FBI)
- Intelligence and Security Command (INSCOM)
- National Counterterrorism Center (NCTC)
- National Geospatial-Intelligence Agency (NGA)
- National Ground Intelligence Center (NGIC)
- National Reconnaissance Office (NRO)
- National Security Agency (NSA)
- National Virtual Translation Center (NVTC)
- Open Source Center (OSC)
- Special Operations Command (SOCOM)
- USA Custom Border Protection (CBP)
- Many Intel Agencies

Where Does NLP Help

- **Data Triage for Analysts & Investigators**
- **Screening at borders**
- **Enrich forensics applications**
- **OSINT**
- **Improve existing applications**

AI / NLP CAN BE AN OBSTACLE IN THE WAY OF TERRORISTS



STANDARD NAME MATCHING FAILS!



Rules-Based Matching



Тамерлан
Царнаев



**Boston Marathon
Bombing**
[~300 Casualties]

INCOMPLETE IDENTITIES

Kadiza Sultana



no issues

Passport



no issues

Home Office



TRAVELLING TOGETHER



Flight Manifest



Shamima Begum



Amira Abase

Image Sources:
- Tweet: [ISD Global](#)

INCOMPLETE IDENTITIES

Kadiza Sultana



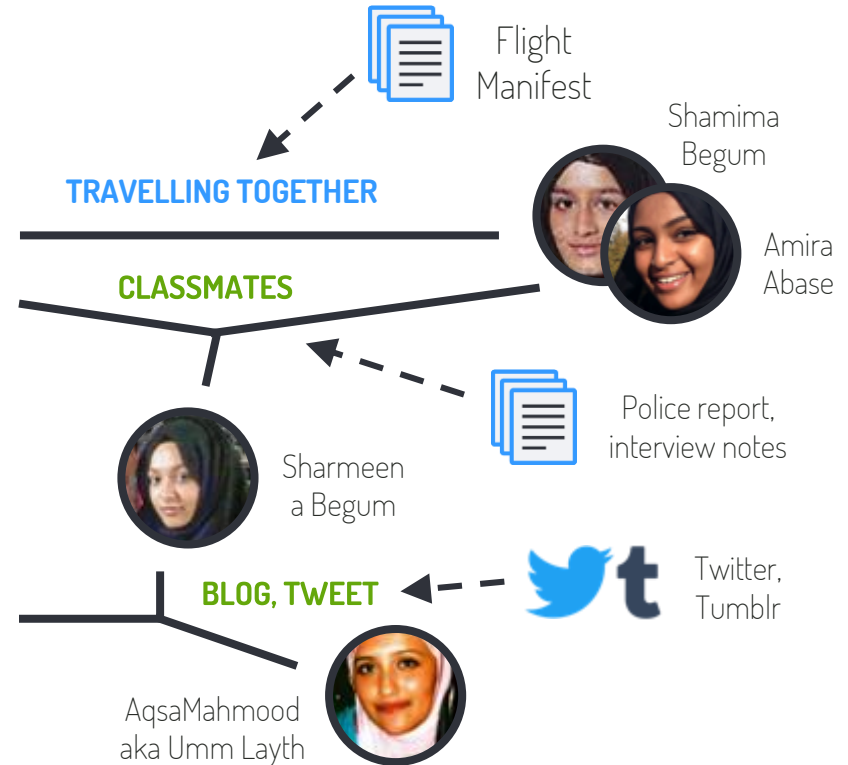
no issues

Passport



no issues

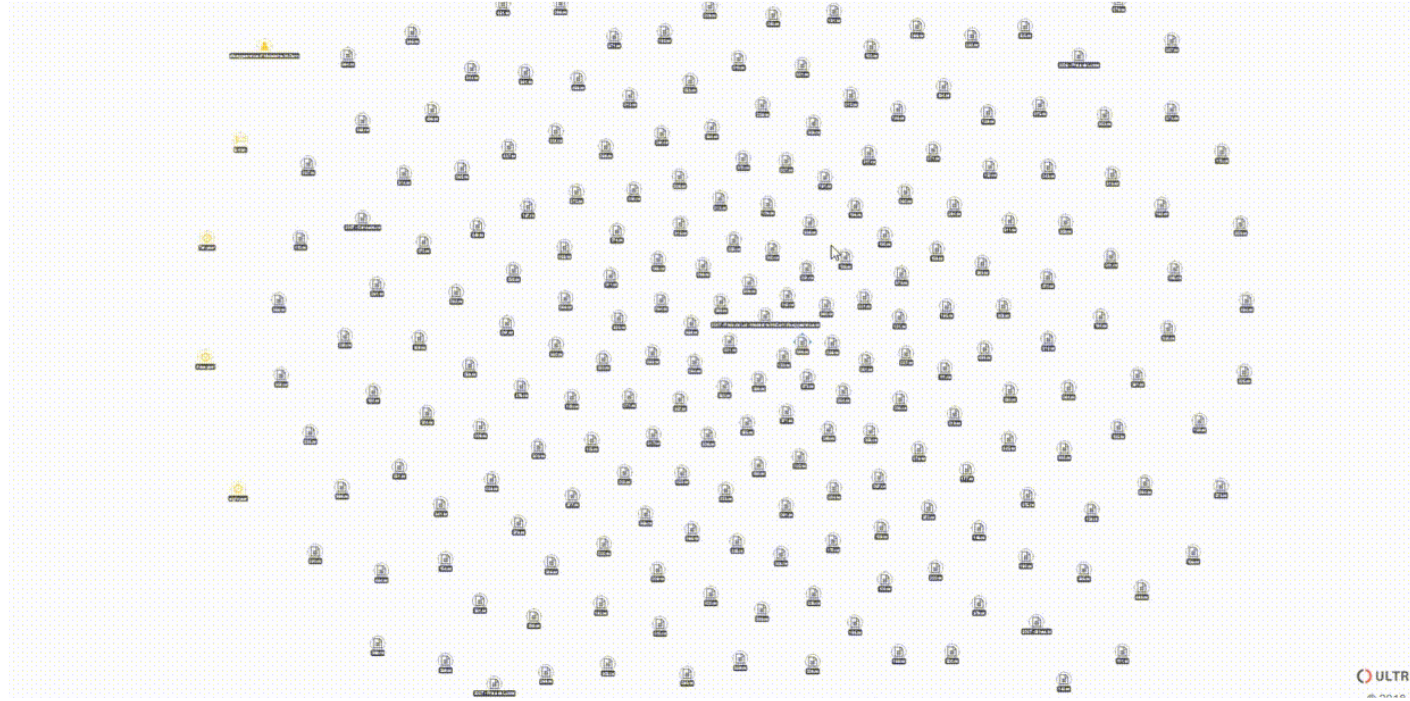
Home Office



LAW ENFORCEMENT - CONNECTING CRIMES



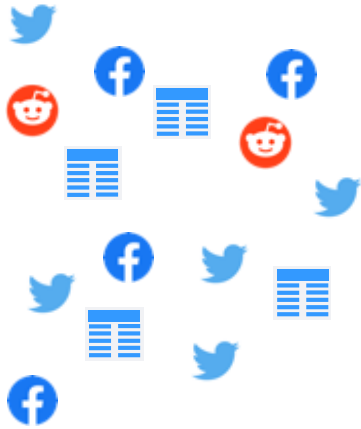
LAW ENFORCEMENT - CONNECTING CRIMES



PREDICTING CIVIL UNREST

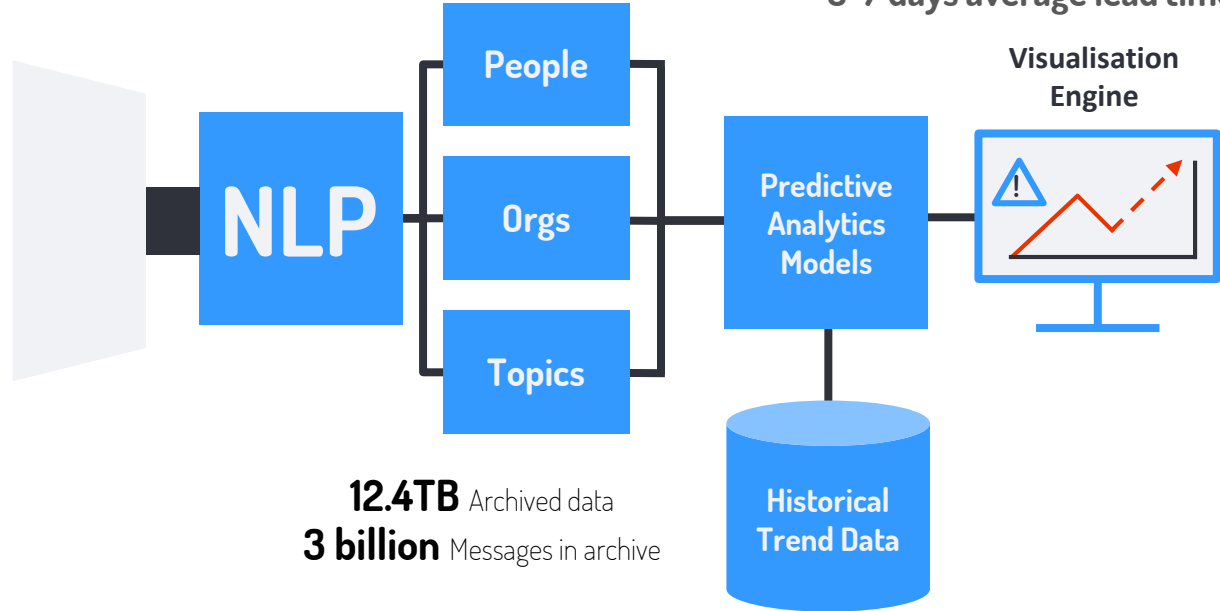
2,000 /sec

Message throughput



20GB

Daily ingest



- 40 – 50 Alerts per day
- 80% – 90% Accurate forecasts
- 5-7 days average lead time

12.4TB Archived data
3 billion Messages in archive

PREDICTING CIVIL UNREST



Improving existing applications and processes

review File Edit View Go Tools Window Help

Screen Shot 2017-09-08 at 9:08:29 PM (1).png

NOW!

Home Index Search Taxonomy Projects Admin

Search

Document list Entity Types Entities Categories **Graph** Relationships Timeline

Person (5301)

- Tony Blair (198)
- Michael Howard (190)
- Gordon Brown (86)
- Oliver (87)
- Charles Kennedy (82)
- Alan Milburn (26)
- Andy Robertson (28)
- David Blunkett (26)
- Jack Straw (26)
- Charles Clarke (27)

Show all >

Organisation (7032)

Location (8190)

- US (549)
- UK (312)
- London (260)
- Britain (238)
- Europe (232)
- England (204)
- France (171)
- Wales (131)
- Scotland (118)
- Ireland (108)

Show all >

ISO Title (4682)

Product (2031)

Temporal: date (2076)

Temporal: time (2685)

Nationality (2315)

- British (202)

Graph visualization showing entities (Person, Location, Product, etc.) and their relationships. The graph is a complex network of nodes and edges, with nodes representing entities and edges representing relationships. The nodes are color-coded and labeled with their respective entity types and names. The graph is displayed in a circular layout, with nodes arranged around a central point and connected by lines. The edges are labeled with relationship types, such as 'is a', 'has a', 'is related to', etc. The graph is a visual representation of the data stored in the application, allowing users to explore the relationships between different entities.

Best Entity Match:

Anwar al-Awlaki

Yemeni-American Imam

Possible Watchlist Match

Name: Anwar AL-AULAKI (YEMENI)
ID: ANWALAKI@Gmail_Co
Score: 0.904

Filter by Anwar al-Awlaki