HOW NLP SAVES THE WORLD!

From Stopping Terrorist Attacks to Spotting Fake News

Chris Brown & Declan Trezise

BASIS TECHNOLOGY

Unstructured Data Workflow

COLLECT

80% of data is UNSTRUCTURED V-2-T

EXTRACT

1) Natural Language Processing Extracts Facts

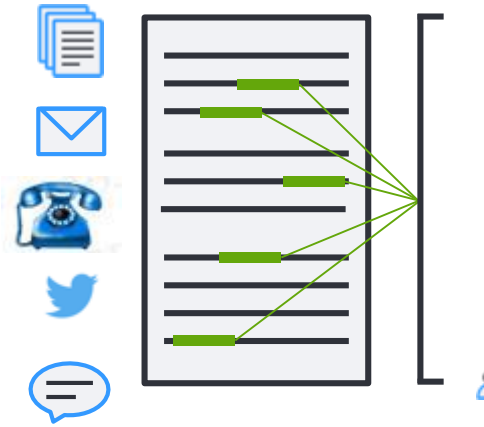
2) Scored for confidence & relevance

COMBINE

Join Processed and Structured Data into Knowledge Graph

ANALYSE

Mine Graph For Patterns & Changes



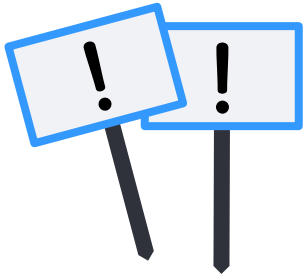
- People
- Organisations
- Locations
- Relationships

- Alerting
- Searching
- Anomaly Detection
- Reporting

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

Tamerlan Tsarnayev → ← **Tamerlan Tsarnaev** → **NO MATCH X** → **Boston Marathon Bombing** [~300 Casualties]

Тамерлан Царнаев

INCOMPLETE IDENTITIES

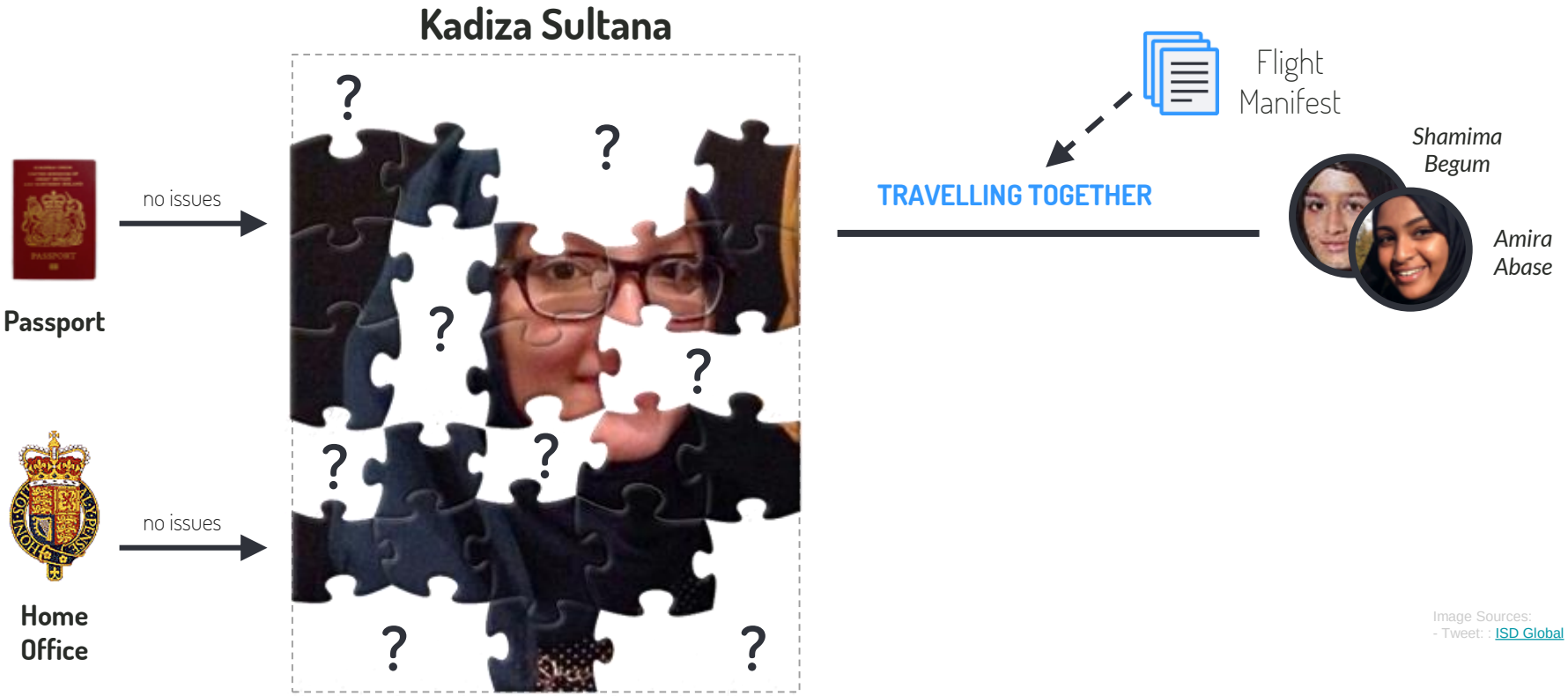
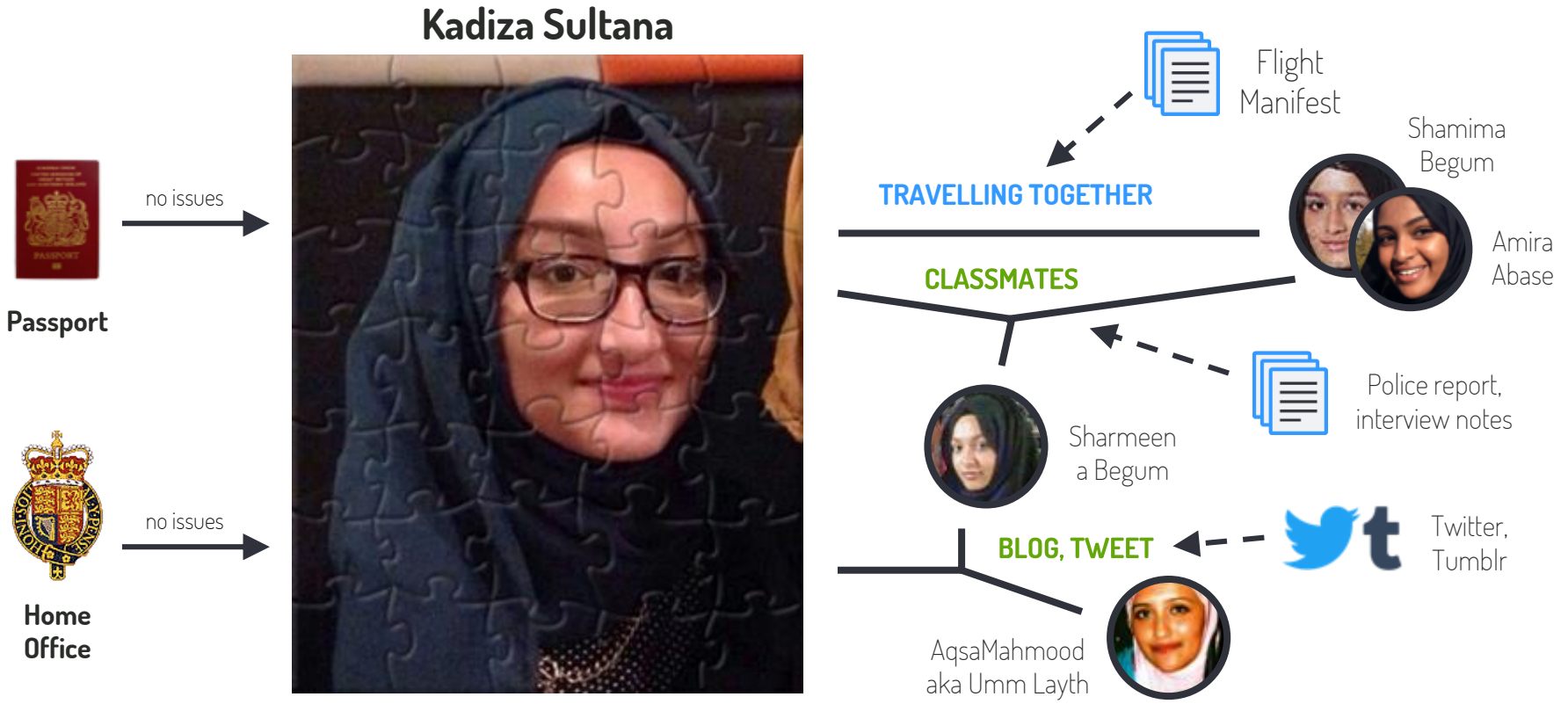


Image Sources:
- Tweet: : ISD Global

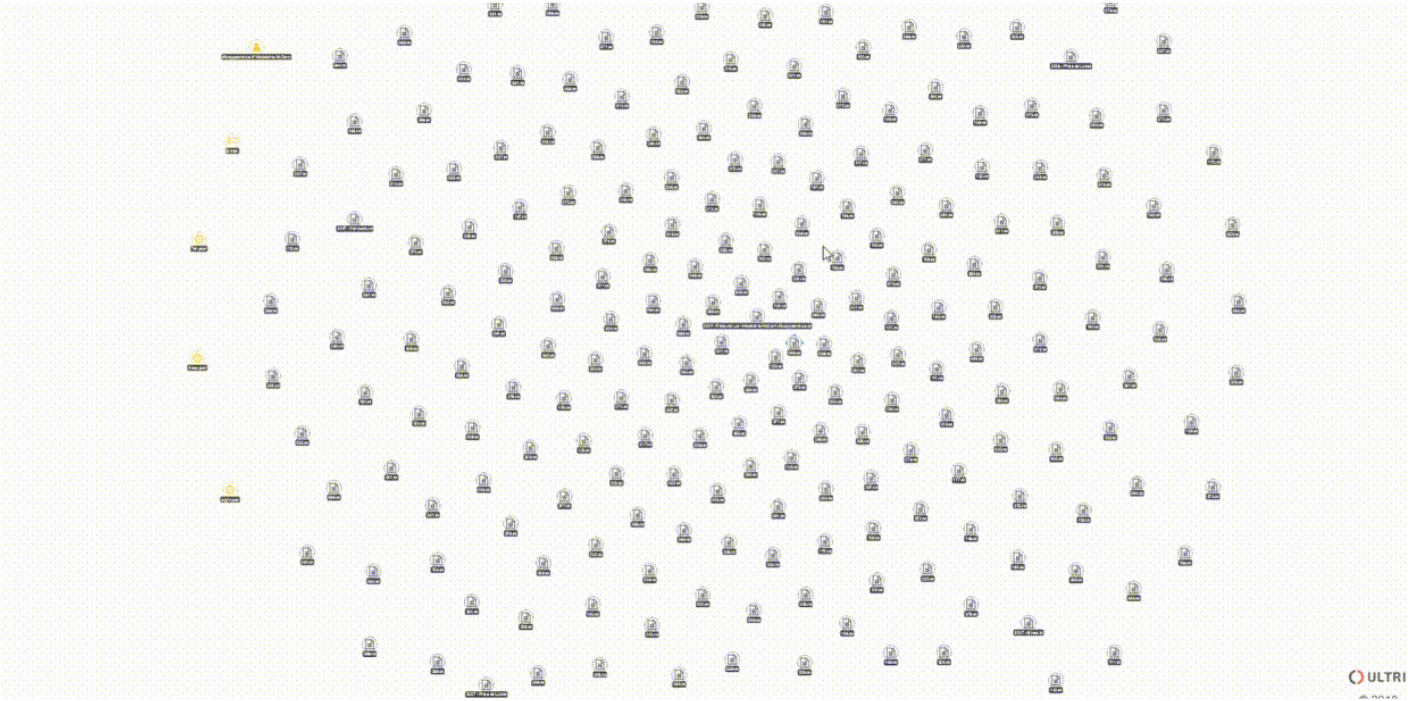
INCOMPLETE IDENTITIES



LAW ENFORCEMENT – CONNECTING CRIMES

Image source: https://www.thesun.co.uk/archives/news/291020/secret-attacks-before-maddie-kidnap/

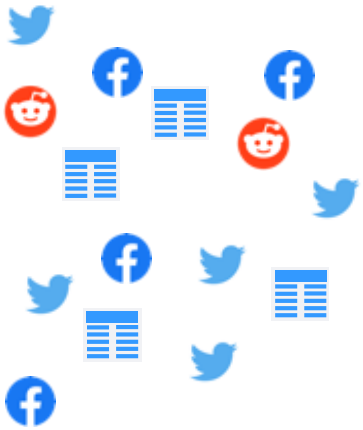
LAW ENFORCEMENT – CONNECTING CRIMES



PREDICTING CIVIL UNREST

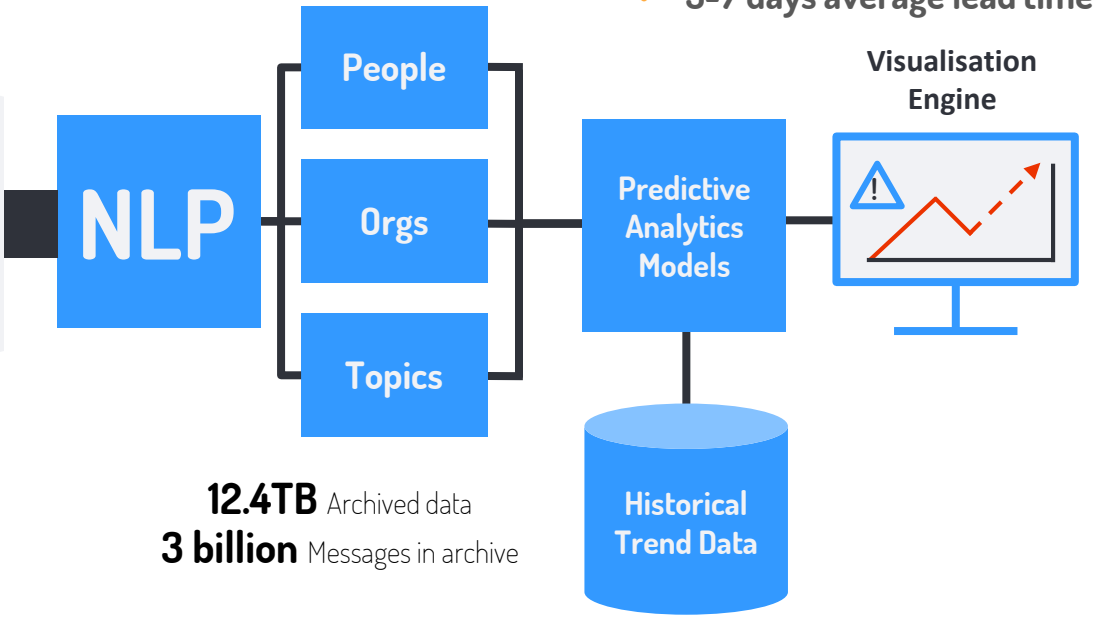
2,000 /sec

Message throughput



20GB

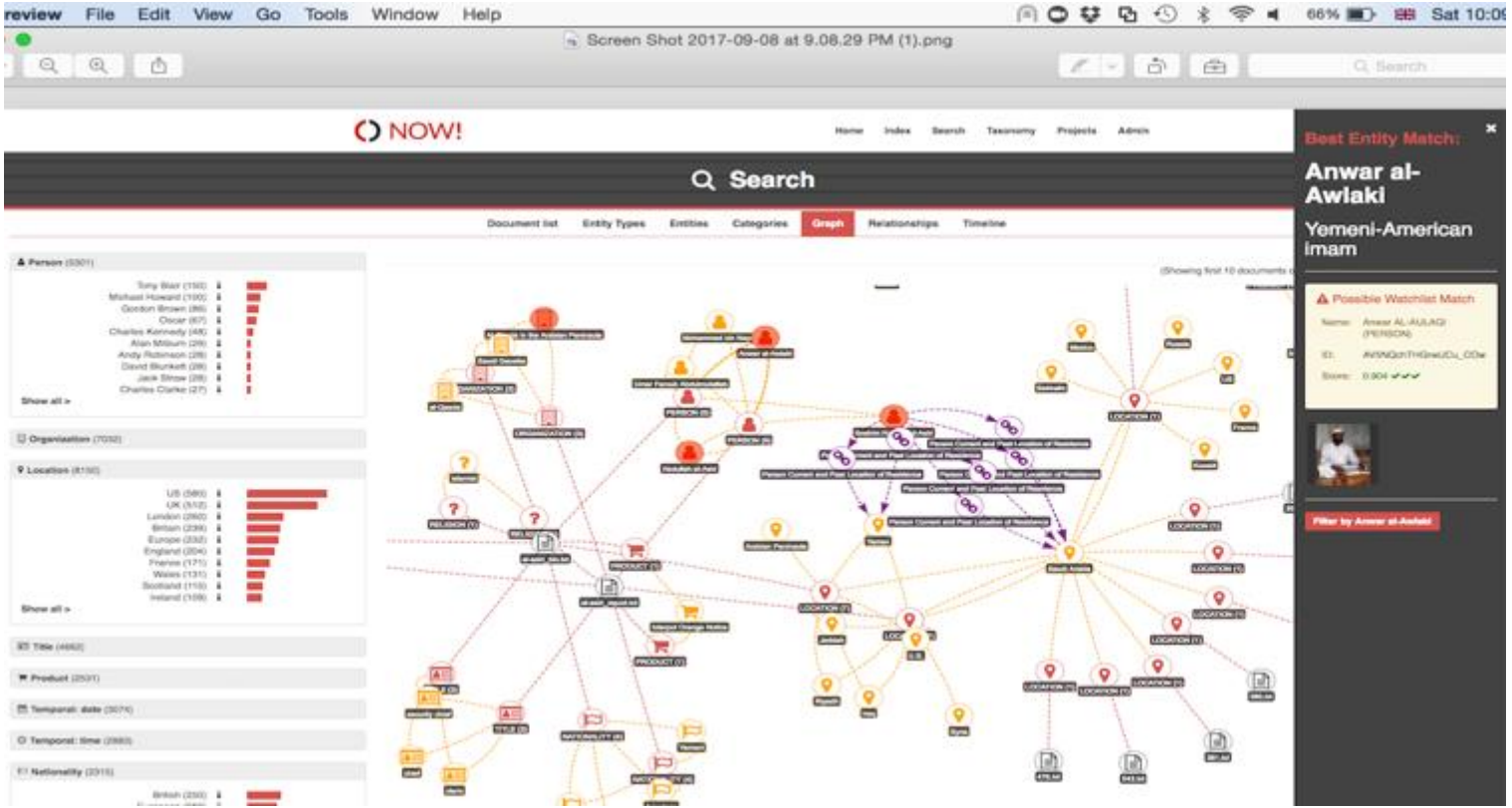
Daily ingest



12.4TB Archived data

3 billion Messages in archive

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

PREDICTING CIVIL UNREST

Improving existing applications and processes